DOCUMENT RESUME

ED 342 244 FL 020 104

AUTHOR Wylie, Elaine, Ed.
TITLE Assessment of Proficiency in Japanese as a Foreign Language. Conference Proceedings of an International Working Group Sponsored by the Asian Studies Council (Canberra, Australia, June 12-14, 1990).
INSTITUTION Asian Studies Council, Canberra (Australia).
REPORT NO ISBN-086-857-333-7
PUB DATE Jun 90
NOTE 144p.
PUB TYPE Collected Works - Conference Proceedings (021)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Air Transportation; Daily Living Skills; Employer Attitudes; Foreign Countries; Hotels; *Japanese; *Language Proficiency; *Language Tests; Oral Language; *Second Languages; Testing; *Tourism; Uncommonly Taught Languages
IDENTIFIERS *Australia

ABSTRACT

The proceedings of a working group conference on proficiency testing of Japanese as a second language contain a brief background paper distributed to conference invitees, a list of items included in the pre-conference portfolio, an advance organizer of potential discussion topics, a 77-item annotated list of bibliographies on second language testing, an additional unannotated 110-item bibliography, a brief list of relevant tests and scales, a conference schedule, a list of participants, and the following papers: "Towards the Development of Proficiency and Other Tests in Japanese as a Foreign Language in Australia" (D. E. Ingram); "Tourism Industry Needs and Priorities" (B. Kinnaird); "Employer Expectations" (R. McLeod); "Assessment of Proficiency Levels--Japanese Language: Relevance to the Hotel Industry" (J. K. Lambert); "Proficiency Levels for Japanese Speakers--Qantas Perspective" (S. O'Donnell); "A Brief Summary of the Interim Reports of Syllabus for the Japanese Language Proficiency Test (1989-90)" (K. Okuda); "Proficiency Testing" (T. Whalley); "Foreign Language Proficiency Assessment in the U.S." (D. V. Whipple and H. Kataoka); "Language Planning and the Testing of Japanese in Australia" (J. V. Neustupny); and "Designing Instruments To Measure Language Proficiency" (A. Davies and A. Brown). An outline of conference activities called "Daily Tasks and Outcomes," and a section describing specific project proposals resulting from the conference, with suggested timelines, staffing, costs, and funding sources, are appended. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED342244

ASSESSMENT OF PROFICIENCY IN JAPANESE AS A FOREIGN LANGUAGE

Conference Proceedings

of an International Working Group

sponsored by the Asian Studies Council

Canberra

12 - 14 June, 1990

edited by

Elaine Wylie

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ingram D.E

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL020 104

Asian Studies Council

BEST COPY AVAILABLE

ASSESSMENT OF PROFICIENCY
IN JAPANESE AS A FOREIGN LANGUAGE

Conference Proceedings

of an International Working Group

sponsored by the Asian Studies Council

Canberra

12 - 14 June, 1990

edited by

Elaine Wylie

© Asian Studies Council, Canberra, 1990

ISBN 086 857 333 7

Printed by Griffith University Reprographics, Brisbane, Australia

TABLE OF CONTENTS

		page
1.	Preface - Background Paper *Asian Studies Council*	1
2.	List of Items included in Pre-Conference Portfolio	3
3.	Advance Organizer *D E Ingram*	4
4.	Bibliographies *D E Ingram*	6
5.	Program	29
6.	List of Participants	31
7.	Overview Paper: Towards the Development of Proficiency and Other Tests in Japanese as a Foreign Language in Australia *D E Ingram*	34
8.	Tourism Industry Needs and Priorities *Bob Kinnaird*	52
9.	Employer Expectations *Roslyn McLeod*	58
10.	Assessment of Proficiency Levels - Japanese Language: Relevance to the Hotel Industry *J K Lambert*	62
11.	Proficiency Levels for Japanese Speakers - Qantas Perspective *Sharon O'Donnell*	66
12.	A Brief Summary of the Interim Reports of Syllabus for the Japanese Language Proficiency Test (1989-90) *Kunio Okuda*	70
13.	Proficiency Testing *Tom Whalley*	73
14.	Foreign Language Proficiency Assessment in the U.S. *David V Hiple and Hiroko Kataoka*	77
15.	Language Planning and the Testing of Japanese in Australia *J V Neustupny*	79
16.	Designing Instruments to Measure Language Proficiency *Alan Davies and Annie Brown*	94
17.	Daily Tasks and Outcomes	104
18.	Resultant Project Proposals	112

PREFACE

[This paper was originally distributed by the Asian Studies Council to invitees to the Working Group]

Background

Recent reports on language policy and studies in Australia have highlighted the need to develop language proficiency testing in second or foreign languages. An important recommendation of the Report on the National Policy on Languages Policy on Languages (1986) was the establishment of a national languages testing co-ordinating unit. That recommendation was endorsed by government and implemented within the Languages Institute of Australia (LIA) The Report of the Inquiry into the Teaching of Asian Studies and Languages in Higher Education by Professor John Ingleson (January 1989) has highlighted the need to develop language proficiency tests which build on recent language learning guidelines and complement the NAATI interpreting/translating levels of proficiency used in Australia. The report proposed that yardsticks for the measurement of various levels of communicative skills should be established in consultation with tertiary language teachers and relevant language associations. These yardsticks would then be applied against the Australian situation to ascertain where changes in emphasis needed to be made.

No proficiency scales have been established for Japanese language learning in Australia. The Japan Foundation administers a written examination of student proficiency, however, the test has limited applicability to a number of contexts. The test does not assess spoken proficiency and is focused on grammar, vocabulary and kanji reading competence. As such the test does not coincide with current approaches to communicative language learning and even the needs of academic institutions in Japan in regard to listening and oral proficiency levels. If found practicable cross referencing between the new scales and this test would be mutually beneficial. Employers in the tourism and the hospitality industry have a particular need to assess proficiency in speaking and listening.

The existence of acceptable proficiency scales reflecting the development of proficiency in Japanese as a second language could make important contributions to the design of Japanese syllabus, teaching materials and classroom practices. The Working Group will need to consider the breadth and focus of any subsequent test development projects for adults, adolescents and younger learners. In summary there is a need to produce proficiency scales and design a battery of tests which will test proficiency in a broad range of contexts. One option for example might be for the test to function in a similar way to the International English Language Testing System which tests all four language skills (reading, writing, listening and speaking) and is taken by non-native English speakers intending to study or work in an English-speaking environment.

The development of proficiency scales and associated tests would be carried out in close consultation with the LIA. One of the main functions of the LIA is to develop proficiency measurement instruments in English and in other languages.

Objective of the Working Group meeting

The meeting would be the starting point from which to develop the means of assessing practical proficiency in Japanese language in Australia for adults, adolescents and younger learners. A number of development tasks will be required in the project including the production of proficiency scales and associated test materials. Direct testing, interviewing and other tests might be involved depending on the precise purposes determined for the test battery to be developed.

Whatever approach to testing is used, proficiency scales will be required.

Expected outcomes from the meeting

An understanding of the field of language proficiency, available proficiency scales in the Japanese language, and how to proceed to construct appropriate scales.

A detailed specification of the range of needs for testing language proficiency and the constraining factors on test development.

Production of outlines of research projects required to develop assessment tools.

A statement of future structures and strategies (including the role of the LIA) with financial implications, as well as implications for other languages. There may also be scope for involvement of other language organisations in the region.

LIST OF ITEMS INCLUDED IN PRE-CONFERENCE PORTFOLIO

Advance Organizer

Bibliographies

Clark, John L. D. 1986. *Handbook for the Development of Tape-Mediated, ACTFL/ILR Scale-based Tests of Speaking Proficiency in the Less Commonly Taught Languages*. Washington, D.C.: Center for Applied Linguistics.

Ingram, D. E. 1990. "The International English Language Testing System (IELTS): Its Nature and Development". Paper to the 1990 RELC Seminar, Regional Language Centre, Singapore, 9-12 April, 1990. (mimeograph)

Ingram, D. E. 1990a. "The International English Language Testing System (IELTS): The Speaking Test". Paper to the 1990 RELC Seminar, Regional Language Centre, Singapore, 9-12 April, 1990. (mimeograph)

Ingram, D. E. 1985. "Assessing Proficiency: An Overview on Some Aspects of Testing". In Hyltenstam and Pienemann 1985: 215-276.

Ingram, D. E. and Elaine Wylie. 1989. "Developing Proficiency Scales for Communicative Assessment". An abridged version of a paper presented at the National Assessment Consultation for the National Assesssment Framework for Languages at Senior Secondary Level, Sydney, 5 December, 1989 (mimeograph).

Ingram, D. E. and Elaine Wylie. 1979/1985. *The Australian Second Language Proficiency Ratings*. (mimeograph)

Ingram, D. E. and Elaine Wylie. 1982. *The Australian Second Language Proficiency Ratings (Version for Japanese as a Foreign Language)*. (mimeograph) [NB. Early draft only - not released for use or citing].

Morrow, Keith. 1990. "Evaluating Communicative Tests". Paper to the 1990 RELC Seminar, Regional Language Centre, Singapore, 9-12 April, 1990. (mimeograph)

Stansfield, Charles W. 1989. *Simulated Oral Proficiency Interviews. ERIC Digest*, December, 1989. Washington, D.C.: Center for Applied Linguistics.

Thompson, Irene, Richard T. Thompson and David Hiple. 1987. "Adapting the ACTFL/ETS Guidelines to the Less Commonly Taught Languages". In Stansfield and Harman 1987, Chapter 4.

Thompson, Richard T. and Dora E. Johnson. 1988. *ERIC Digest: Proficiency Testing in the Less Commonly Taught Languages*. Washington, D.C.: Center for Applied Linguistics.

ADVANCE ORGANIZER

D. E. INGRAM

The purpose of this sheet is to suggest things to which participants at the forthcoming conference (12-14 June, 1990) may wish to give some thought prior to the conference.

1. A number of papers are enclosed that raise i sue pertinent to the sort of issues to be considered. If anyone is not aware of the sorts of issues raised in these papers, it would be useful to peruse the enclosed papers in advance. In particular, the bibliographies were compiled to give an overview of the field of proficiency and its assessment especially insofar as they might be pertinent to the conference and any subsequent test development projects.

2. What are the purposes for which we need tests relevant to the learning and teaching of Japanese in Australia? Consider purposes that might exist both within education and outside in industry, the community, etc. Consider, amongst many others, such needs as measuring proficiency, course attainment, training standards for teachers, vocational needs, salary loadings, staff recruitment, matching employees to tasks, entry to courses in Australia and overseas (including Japan), streaming, selection of staff for further language training, rational course planning, etc.

3. What are the practical and situational constraints on the sorts of tests that might be devised for these purposes?

4. Is it only tests of proficiency that are required or are other types of tests also desirable? If so, for what purposes and with what constraints on their use? Consider, amongst many others, such needs as diagnosis of learning difficulties, aptitude and the many related issues, course attainment, vocational needs, matching employees to tasks, entry to courses in Australia and overseas (including Japan), streaming, selection of staff for language training, attitudes, cultural understanding, etc.

5. Is there need for tests of Japanese for Special Purposes (including Japanese for Academic Purposes)? If so, what are the "special purpose" contexts for which there is greatest need (e.g., airline booking clerks, scuba diver instructors, banking, automotive engineering, travel agents, academic purposes, etc. etc.)?

6. What tests currently exist, what are their limitations, and can they be modified in some way to be more suitable? Alternatively, are new instruments required and, if so, how would they relate to existing tests?

7. Are proficiency rating scales desirable and, if so, what form should they take? Are any existing scales suitable as they currently stand? If not, what modifications are required or how should we go about developing alternative scales? What special issues need to be taken into account during such a project?

8. In order to meet the educational needs, do we require one scale for all age groups or do we need to consider developing separate scales for children and adults?

9. What are the special implications of the Japanese orthography?

10. Is there need for a test or test battery that provides a single figure score or is a profile result that distinguishes various skills desirable (e.g., many scales and tests distinguish speaking, listening, reading and writing)?

11. If proficiency scales are to be produced, how are they to be used and applied (e.g., through a live interview, through a recorded interview such as a Simulated Oral Proficiency Interview, or through other test types)? What are the test development implications?

12. Is a test battery using a variety of test types (cf. IELTS) needed to allow proficiency to be measured even when trained interviewers with high levels of proficiency are not available?

13. Is there need for approaches to be developed for self-assessment in Japanese?

14. What input to syllabus design, materials development, and classroom teaching would tests and/or proficiency scales have at the different educational levels? How would they relate to the national curriculum materials being developed?

15. How might new tests and proficiency scales contribute to making the teaching of Japanese in schools and universities more effective and efficient?

16. What needs are there in industry and commerce for tests of proficiency in Japanese and/or for persons able to advise on requirements for language skills and on the skills of staff or potential staff?

17. Is there need for accredited language assessors?

18. Are the needs in Australia sufficiently similar to the needs elsewhere that the tests and scales to be developed would be useful elsewhere? If so, is cooperation in the development of suitable tests and scales a practical possibility?

19. Specifically what test development projects are required to meet the needs, how do they fit together to make a comprehensive and prioritized schedule for the future, what sources of funding are there, and how would what is proposed here relate to existing organizations such as the National Language Institute of Australia, the Japan Foundation, etc?

20. To what extent are the issues raised here and any consequent test development requirements relevant to other languages? Are separate test development projects required or can there be rationalization of some sort?

BIBLIOGRAPHIES

compiled by

D. E. INGRAM

NOTE: Searches were conducted both through Australian libraries and through the Center for Applied Linguistics in Washington, D.C.. Very few items were found linking Japanese and language testing. The following very selective bibliographies are not intended to be comprehensive (the literature is vast) but they indicate some items in the general area of language proficiency and its assessment that may be useful as work is commenced on particular projects. Additional, more specific searches or a re-reading of the ones made for this conference should be carried out as each particular project is commenced. The bibliographies provided are in four sections: A. Ready-made bibliographies from various sources; B. an Annotated Bibliography; C. a General Bibliography; D. Relevant Tests, Scales, etc.

A. READY-MADE BIBLIOGRAPHIES

1. ERIC Clearinghouse on Languages and Linguistics. 1990. Computer Search No. 917: *Oral Proficiency Testing*. Washington, D.C.: Center for Applied Linguistics.

2. Regional Language Centre. 1990. *Specialized Bibliography 48: Language Testing and Language Programme Evaluation*. Singapore: RELC.

3. ELTS Revision Project Team. 1989. "Bibliodat", bibliography compiled in the course of the ELTS Revision Project and development of the International English Language Testing System. Cambridge: UCLES. (mimeograph)

4. Galloway, Vicki, Charles W. Stansfield, and Lynn E. Thompson. 1987. "Topical Bibliography of Proficiency-Related Issues". In Stansfield and Harman 1987, Chapter 6.

B. ANNOTATED BIBLIOGRAPHY

Alderson, J. Charles, Karl J. Krahnke and Charles W. Stansfield. 1987. *Reviews of English Languge Proficiency Tests*. Washington, D.C.: TESOL.

This volume contains reviews of some 46 tests of English used for a variety of purposes including proficiency assessment. Though the views expressed are those of the reviewers (which, in the opinion of the volume's editors, sometimes significantly qualifies the worth of some reviews and their insightfulness), it is a useful reference on the range of test types available and the strengths and weaknesses to be found in tests.

Allen, J.P.B. and Alan Davies. (eds.). 1977. *Testing and Experimental Methods, The Edinburgh Course in Applied Linguistics, Vol. 4.* London: O.U.P.

This is a standard reference in language testing though with little on the concept of proficiency and its assessment. It has chapters on basic concepts in testing, the construction of language tests, the design and interpretation of experiments, and statistical procedures.

Bachman, Lyle F. and Adrian S. Palmer. 1981. "The Construct Validation of the FSI Oral Proficiency Interview". In *Language Learning*, Vol. 31, No. 1, 1981, pp. 67-86.

The authors examine the construct validity of the FSI oral interview using a multitrait-multimethod matrix comprising six measures. The results indicate both convergent and discriminant validity for the FSI interview and the authors conclude that the results provide strong evidence against a model of unitary language competence and for a partially divisible language competence.

Barnwell, David. 1986. "Who is to Judge how well Others Speak? An Experiment with the ACTFL/ETS Oral Proficiency Scale", paper to the Eastern States Conference on Linguistics. ED296589 FL017471.

Seven teachers of Spanish were given brief formal training on the use of the ACTFL Scale and then they evaluated recorded interviews of learners of Spanish at different proficiency levels. A high degree of agreement amongst the raters was found despite their minimal training.

Biersteker, Ann. 1986. "A Language Proficiency Testing Model for Business Needs". In *Proceedings of the Fifth Annual Eastern Michigan University Conference on Languages for Business and the Professions.* ED295447 FL016621.

The ACTFL OPI was adapted to focus on business contexts with proficiency profiles adapted accordingly. These were found to be more specific in identifying skills and needs for business purposes and suggested proficiency guidelines are proposed.

Byrnes, Heidi and Michael Canale. 1987. *Defining and Developing Proficiency: Guidelines, Implementations, and Concepts.* Lincolnwood, Ill.: National Textbook Company.

Three of the five chapters in this volume concern the ACTFL Guidelines while the fourth concerns the application of proficiency levels to classifying texts and the last discusses second language acquisition and the implications of proficiency for course design and teaching methods.

Canale, Michael and Merrill Swain. 1980. "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing". In *Applied Linguistics*, Vol. 1, No. 1, 1980, pp. 1-47.

This paper examines the concept of communicative competence looking at some of the theories that have been produced. The authors propose a theoretical framework for communicative competence and examine its implications for foreign language teaching and testing.

Carroll, Brendan J. 1980. *Testing Communicative Performance*. Oxford: Pergamon.

Carroll discusses the assessment of communicative performance and the design, development, and uses of communicative tests. The last chapter reviews some relevant literature in the area of communicative testing. Lengthy appendices include a needs profile, "information banks" for test development (actually specifications for an RSA test), and sample rating scales.

Clark, John L.D. 1988. "Validation of a Tape-Mediated ACTFL/ILR/Scale-based Test of Chinese Speaking Proficiency". In *Language Testing*, Vol. 5, No. 2, December 1988, pp. 187-205.

This paper reports a study in which students' performance on a simulated oral proficiency test is compared with their performance on a live interview. Results were similar though candidates found the simulated interview more difficult.

Clark, John L. D. 1986. *Handbook for the Development of Tape-Mediated, ACTFL/ILR Scale-based Tests of Speaking Proficiency in the Less Commonly Taught Languges*. Washington, D.C.: Center for Applied Linguistics.

This small booklet documents the development of a tape-based test for Chinese. The test is a "proficiency-oriented semi-direct test" which approximates to the interview situation but the stimuli and the candidate's responses are recorded and the candidate's performance is rated using the ACTFL/ILR Proficiency Guidelines. The procedures and issues raised are relevant to other languages and may be especially useful where live interviewing is not possible.

Clark, John L. D. (ed.). 1978. *Direct Testing of Speaking Proficiency: Theory and Application*. Princeton, NJ: Educational Testing Service.

This volume contains eighteen papers from a two-day conference on the assessment of speaking proficiency. Most interest focussed around the development and use of the FSI Scale and the associated interview procedures. Other topics included the use of the FSI Scale and interview in different contexts for different purposes including diagnostic evaluation, teacher certification, and High School attainment.

Clark, John L. D. 1972. *Foreign-Language Testing: Theory and Practice.* Philadelphia, Pa.: Center for Curriculum Development.

This volume considers different types of testing including prognostic testing and tests of achievement (all four macroskills), knowledge (including culture and literature), and proficiency (with special reference to the FSI Scale and interview). It also discusses the range and roles of published tests.

Clark, John L. D. and Ray T. Clifford. 1988. "The FSI/ILR/ACTFL Proficiency Scales and Testing Techniques: Development, Current Status, and Needed Research". In *Studies in Second Language Acquisition*, Vol. 10, No. 2, June 1988, pp. 129-147.

Describes the rationale, history, and current state of the proficiency movement in the U.S. and identifies areas requiring further research and development.

Clark, John L. D. and Eleanor H. Jorden. 1984. *A Study of Language Attrition in Former U.S. Students of Japanese and Implications for Design of Curriculum and Teaching Materials: Final Project Report.* Washington, D.C.: Center for Applied Linguistics.

This publication reports a long-term study on the attrition of skills in Japanese amongst university students of Japanese. The period of attrition differed. Initial proficiency (measured using the Japanese Proficiency Test and a self-assessment schedule) seemed to have little influence on the extent of attrition, decreases in proficiency were inversely related to the amount of use made of the language, attrition seemed to be skill-specific, and re-acquisition of skills seemed to be faster using familiar rather than unfamiliar materials. A number of case studies are provided and confirmed the group data analyses.

Cziko, Gary A. 1981. "Psychometric and Edumetric Approaches to Language Testing: Implications and Applications". In *Applied Linguistics*, Vol. 2, No. 1, Spring 1981, pp. 27-44.

The essential differences between psychometric and edumetric tests in education are outlined, implications drawn for language testing, discrete-point and integrative approaches reviewed, and the construction and use of edumetric integrative tests are described.

Davies, Alan. 1984. "Computer-Assisted Language Testing". In *Calico Journal*, Vol. 1, No. 5, June 1984, pp. 41-42 and 48.

This short paper reviews the state of computer-assisted language testing in the first part of the 1980s and identifies its strengths and limitations. It identifies three areas of particular interest: testing for self-access purposes, cloze and reading-speed tests, and diagnostic testing.

Doyle, Michael S. 1988. "Business Spanish: Role Playing for Oral Proficiency". In *Proceedings of the Seventh Annual Eastern Michigan University Conference on Language for Business and the Professions.* ED304938 FL017895.

This reports the use of ACTFL OPI situation cards in special purpose areas of the language for both measurement and teaching purposes.

Dugan, J. Sanford. 1988. "Standardized Tests as an Alternative to the Oral Interview". In *Proceedings of the Seventh Annual Eastern Michigan University Conference on Languages for Business and the Professions.* ED304923 FL017880.

A moderate correlation was found between the MLA listening and speaking tests and the ACTFL/ETS OPI results. Problems in interpreting the MLA results were also noted.

Ellis, Rod. 1984. "Communication Strategies and the Evaluation of Communicative Performance". In *ELT Journal*, Vol. 38, No. 1, January 1984, pp. 39-44.

Ellis proposes that, instead of trying to define proficiency or communicative competence, the tester should focus on "communication strategies" for assessment purposes, i.e., strategies adopted by learners in order to communicate when they lack the necessary language.

Frith, James R. (ed.). 1980. *Measuring Spoken Language Proficiency.* Washington, D.C.: Georgetown University Press.

This small volume contains papers relevant to the measurement of speaking proficiency. It considers factors in speaking proficiency, ways of measuring speaking proficiency using both live and recorded interviews, and examples of the use of oral proficiency ratings.

Green, Peter S. (ed.). 1987. *Communicative Language Testing: A Resource Book for Teacher Trainers.* Strasbourg: Council for Cultural Cooperation.

This book is intended as a resource for teacher trainers who wish to organize inservice programmes on the communicative testing of foreign language competence in High School students. The proposals are organized under six themes including the nature of communication, the content of a communicative testing syllabus, testing communicative competence at a national level, classroom assessment, teacher training in testing communicative competence, and testing communicative performance.

Greenberg, Karen. 1986. "The Development and Validation of the TOEFL Writing Test: A Discussion of TOEFL Research Reports 15 and 19". In *TESOL Quarterly*, Vol. 20, No. 3, 1986, pp. 531-544.

This paper reviews the research conducted prior to the introduction of a writing test in TOEFL, looking especially at TOEFL Research Report No. 15 (*Survey of Academic Writing Tasks Required of Graduate and Undergraduate Foregin Students*) and TOEFL Research Report No. 19 (*Relationship of Admission Test Scores to Writing Performance of Native and Nonnative Speakers of Egnlish*). The paper outlines the writing tasks identified and questions the choice of tasks on which TOEFL focusses. At the same time the author supports the use of direct assessment approaches.

Griffin, Patrick et al. 1988. "An Algorithmic Approach to Prescriptive Assessment in English as a Second Language". In *Language Testing*, Vol. 5, No. 1, June 1988, pp. 1-18.

Describes the development using specially developed algorithms of interview items for use with the ASLPR in the range of 0 to 1+.

Harris, David P. 1969. *Testing English as a Second Language*. New York: McGraw-Hill.

This book pre-dates developments in the concept and testing of proficiency and is included, like the books by Valette and Lado, as a standard reference in more traditional aspects of language testing. It discusses basic concepts such as validity and reliability and testing the elements of the language (grammar, pronunciation, vocabulary, etc) as well as the four macroskills. It has chapters on constructing tests, administering tests, interpreting and using test results, and basic statistics.

Heaton, J. B. 1975/1988. *Writing English Language Tests*. London: Longman.

A general and fairly traditional text on language testing with most emphasis on objective tests. It contains a general discussion of language testing and focusses on testing the elements of the language and the macroskills in six of the ten chapters. Other chapters consider standard test evaluation concepts and procedures and the interpretation of test scores.

Henning, Grant. 1983. "Oral Proficiency Testing: Comparative Validities of Interview, Imitation, and Completion Methods". In *Language Learning*, Vol. 33, No. 3, 1983, pp. 315-332.

Henning compares three oral testing methods: imitation, completion and FSI-type interview. Rankings in terms of overall validity across all indexes used were pronunciation, interview and completion. Regression analysis showed that the interview was more closely related to the grammar skill.

Higgs, Theodore V. (ed.). 1985. *Teaching for Proficiency, the Organizing Principle*. Lincolnwood, Ill.: National Textbook Company.

The main focus of the papers in this volume are on the application of the concept of proficiency and its assessment to language teaching though there are also papers on testing. The background and origins of the ACTFL Proficiency Guidelines are traced and the appendix reproduces the Guidelines and the ILR level definitions for speaking.

Ingram, D. E.. forthcoming. "Assessing Special Purpose Proficiency", Chapter One of Ingram, D. E.. in preparation. *Direct Proficiency Assessment*.

This paper discusses the nature of special purpose proficiency and its development and considers the relevance to it of direct proficiency assessment using rating scales. It proposes the development of a special purpose version of the Australian Second Language Proficiency Ratings in which the "General Description" column is the same but there is an additional column describing features specific to the special purpose area in question, an Example column and a Comment column.

Ingram, D. E. 1986. "Proficiency in Syllabus Design and Assessment". Paper to the National Conference and Seminar of the New Zealand Association of Language Teachers, Dunedin, 13-15 May, 1986. ED276246 FL015792

This paper discusses the relevance of the concept of proficiency to syllabus design and assessment giving emphasis to identifying the learner's real needs and the precise specification of goals in proficiency terms.

Ingram, D. E. 1984. *Report on the Formal Trialling of the Australian Second Language Proficiency Ratings (ASLPR)*. Canberra: DIEA/AGPS.

This publication reports the results of the formal trials of the ASLPR which examined the reliability of ratings assigned by native and non-native English speaking raters and the relationship between scores on the ASLPR and in pencil-and-paper tests.

James, Charles J. (ed). 1985. *Foreign Language Proficiency in the Classroom and Beyond*. Hastings-on-Hudson, N.Y.: ACTFL.

Proficiency and its meaning for teachers, curriculum developers, researchers, publishers, and business people are discussed dealing with such issues as definition of proficiency, curriculum and materials development, assessment, and applications beyond the classroom.

James, Gregory. 1984. "Development of an Oral Proficiency Component in a Test in English for Academic Purposes". Based on a paper to the 1984 AILA Congress, Brussels. ED256153 FL014895.

A test of oral proficiency in academic contexts is described. The test, part of the TEEP Test, uses a prerecorded test to which the candidate listens and responds preferably in a language laboratory. Candidates demonstrate their proficiency in free speech, oral summary of material presented graphically, asking and answering questions, responding to remarks and situations, and responding to and expressing opinions.

Jones, Randall L. and Spolsky, Bernard. (eds.). 1975. *Testing Language Proficiency*. Washington, D.C.: Center for Applied Linguistics.

These are the proceedings of the 1974 Washington Language Testing Symposium which focussed especially on language proficiency assessment as it relates to the use of foreign languages at work. Its sections include papers on instances where proficiency testing is used and on assessment of proficiency in speaking, listening and reading.

Lado, Robert. 1961. *Language Testing*. London: Longman.

This book pre-dates developments in the concept and testing of proficiency and is included, like the books by Harris and Valette, as a standard reference in more traditional aspects of language testing. It was published at the height of audiolingualism, behaviourism and structural linguistics and is a standard reference for test types of the era. Its first part discusses a theory of foreign language testing and the basic criteria for the evaluation of tests. The second part has chapters focussing in great detail on the testing of elements of the language (e.g., pronunciation, grammar, vocabulary) and the third part focusses on the macroskills. Part Four considers the testing of cross-cultural understanding and attitudes, while Part Five discusses in some detail such concepts as norms, validity, reliability, item analysis and scoring, has a chapter on different purposes for teting (e.g., diagnostic, achievement and aptitude tests), and one on designing experiments in foreign language learning.

Liskin-Gasparro, Judith E. 1984. "Foreign Language Oral Proficiency Testing: An Organizing Principle for Instruction". Paper to the Annual Meeting of the National Council on Measurement in Education, New Orleans. ED263787 FL015412.

This paper refers to programmes in which proficiency scales have been used to modify the nature of the syllabus and how it is taught using the ACTFL/ETS guidelines as an organizing principle.

Magnan, Sally S. 1988. "Grammar and the ACTFL Oral Proficiency Interview: Discussion and Data". In *Modern Language Journal*, Vol. 72, No. 3, Fall 1988, pp. 266-276.

Discusses the role of grammar in the ACTFL Scale and provides data on the grammatical accuracy of 40 interviews at various proficiency levels of French.

Moeller, Paulette. 1988. "Quality Control and French Proficiency". In *Education Canada*, Vol. 28, No. 2, Summer 1988, pp. 4-9.

This paper proposes the use of the ACTFL Scale and interview as quality control measures for students and teachers in Canada.

Munson, Johanna and Mary Harrison. 1990. "Review of the Japanese Language Proficiency Test", MATFL language testing seminar paper, Monterey Institute of International Studies. (mimeograph)

This short paper reviews the Japanese Language Proficiency Test [see **Section D. Relevant Tests, Scales, etc.** below], highlights its limitations as a proficiency test, and points to some construction faults in the items in the version reviewed. It draws attention to the test's failure to measure productive skills.

Oller, John W. and Kyle Perkins (eds.). 1980. *Research in Language Testing*. Rowley, Mass.: Newbury House.

This volume is in seven parts. The first discusses the factors that make up language skill, the next four focus on proficiency and its measurement in each macroskill, the sixth considers the differences between native and non-native performance, and the last discusses measurement of factors that contribute to language learning. The only article specifically relevant to learning Japanese as a foreign language is one that considers the value of the MLAT aptitude test in predicting success in a university Japanese course. Only moderate correlations were found and other factors (e.g., experience in using the language with native speakers etc) were demonstrated to be also important.

Oskarsson, Mats. 1989. "Self-Assessment of Language Proficiency: Rationale and Applications". In *Language Testing*, Vol. 6, No. 1, 1989, pp. 1-13.

This paper provides a justification for self-assessment, indicates the benefits that flow to language learning from incorporating self-assessment in a language programme, and outlines some suitable techniques ranging from progress cards to rating scales to video and audio recordings.

Oskarsson, Mats. 1978. *Approaches to Self-Assesment in Foreign Language Learning*. Oxford: Pergamon.

This document was prepared in the context of the Council of Europe's work on language learning. It seeks to explore ways in which learners may be assisted to realize their own levels of achievement. It outlines some possible forms of self-assessment (both formal and informal), instances of where self-assessment is being developed and used, and sketches some research experiments.

Politzer, Robert L. and Mary McGroarty. 1983. "A Discrete Point Test of Communicative Competence". In *IRAL*, Vol. 21, No. 3, August 1983, pp. 179-191.

Politizer and McGroarty identify three definitions for communicative competence (ability to receive and convey information; appropriateness to topic, place and interlocutor; effective performance of speech acts). Three tests are devised to test each of these and the results compared wth a grammar test. Linguistic competence is shown to be a condition for communicative competence but students can have the same level in the former and differ widely on the latter.

Raffaldini, Tina. 1988. "The Use of Situation Tests as Measures of Communicative Ability". In *Studies in Second Language Acquisition*, Vol. 10, No. 2, June 1988, pp. 197-216.

Distinguishes the oral proficiency interview and what it measures from communicative ability and proposes two situation tests as alternatives.

Read, John A. S.. 1981. *Directions in Language Testing*. Singapore: RELC.

This collection of papers is from the 1980 RELC Seminar and includes papers on validating language tests, assessing oral proficiency, measuring communicative competence, cloze procedure, self-assessment, and the relationship between testing and teaching.

Rivera, Charlene. (ed.). 1984. *Language Proficiency and Academic Achievement*. Clevedon: Multilingual Matters.

These papers were presented at the Language Proficiency Symposium in Warrenton, Virginia, in 1981. They discuss theoretical frameworks for the concept of proficiency, social and cultural aspects of proficiency, and the need for research to clarify the concept.

Rivera, Charlene. (ed.). 1984a. *Communicative Competence Approaches to Language Proficiency Assessment: Research and Application*. Clevedon: Multilingual Matters.

These papers were presented at the Language Proficiency Assessment Symposium in March 1981 (presumably the same event as for the previous entry). The papers deal with definitions of terms (especially communicative competence), implications for testing, examples of proficiency assessment in different contexts, and psycholinguistic issues.

Savignon, Sandra J. 1985. "Evaluation of Communicative Competence: The ACTFL Provisional Proficiency Guidelines". In *The Modern Language Journal*, Vol. 69, No. 2, 1985, pp. 129-134.

This paper offers a critique of the ACTFL Guidelines, criticizing the inadequacy of the concept of "proficiency" , preferring "communicative competence" but wanting a more adequate model, and asserting that they are unable to provide a basis for syllabus design.

Shohamy, Elana. 1982. "Predicting Speaking Proficiency from Cloze Tests: Theoretical and Practical Considerations for Test Substitution". In *Applied Linguistics*, Vol. 3, No. 3, pp. 161-171.

Shohamy compares results on cloze and oral interviews and concludes that it is better to look at ways of making the oral interview more efficient than to try to substitute other test types (specifically cloze) for it.

Skehan, Peter. 1988. "Language Testing. Part I", "State of the Art Article" in *Language Teaching*, Vol.21, No.4, pp. 211-221.

Skehan, Peter. 1989. "Language Testing. Part II", "State of the Art Article" in *Language Teaching*, Vol.22, No.1, pp. 1-13.

As the title indicates, this two-part article seeks to review current developments in language testing. It considers the nature of proficiency, communicative tests, ESP, achievement testing (including graded objectives), influences on test performance, statistical and analytic techniques, decision-making and test referencing, and public examinations.

Spolsky, Bernard. 1985. "The Limits of Authenticity in Language Testing". In *Language Testing*, Vol. 2, No. 1, 1985, pp. 31-40.

Spolsky argues for the importance of authenticity in test tasks since lack of authenticity weakens the generalizability of the results. All tests are inauthentic and only part of the problem is overcome with authentic-seeming tasks. Real-life observation is suggested as the only solution provided that the observers are sympathetic and trained.

Stansfield, Charles W. and Chip Harman (eds.). 1987. *ACTFL Proficiency Guidelines for the Less Commonly Taught Languages. A Familiarization Project for the Development of Proficiency Guidelines for Less Commonly Taught Languages*. New York/Washington D.C.: ACTFL/Center for Applied LInguistics.

This is a collection of papers on the use of the ACTFL Guidelines for teaching and assesment of "less commonly taught" languages. Papers include discussion of the OPI, the ILR interview, adaptation of the ACTFL Proficiency Guidelines to the less commonly taught languages, materials development, and a bibliography. There is specific discussion of the application of the ACTFL Guidelines to Arabic (a scale is provided), Indonesian, Hindi and African languages.

Stansfield, Charles W. and Dorry Mann Kenyon. 1988. *Development of the Portuguese Speaking Test. Year One Project Report. Development of Semi-Direct Tests of Oral Proficiency in Hausa, Hebrew, Indonesian and Portuguese*. Washington D.C.: Center for Applied Linguistics.

This publication reports the development of a SOPI in Portuguese. Interrater reliabilites were found to be uniform across both the taped and live tests. Most respondents preferred the live to the taped version.

Swain, Merrill. 1984. "Teaching and Testing Communicatively". In *TESL Talk*, Vol. 15, Nos. 1/2, 1984, pp. 7-18.

This paper considers the general principles of communicative language testing and outlines a testing unit in a Grade 9 French course.

Trim, J. (ed.). 1988. *Evaluation and Testing in the Learning and Teaching of Languages for Communication*. Strasbourg: Council for Cultural Cooperation.

This report emanates from the Council of Europe's Project No. 12 "Learning and Teaching Modern Languages for Communication". Its papers deal with a number of issues relevant to communicative testing including assessing tests, personality as a factor in test

development, graded objectives, educational evaluation, self-assessment and automatic data processing.

Underhill, Nic. 1987. *Testing Spoken Language: A Handbook of Oral Testing Techniques*. Cambridge: University Press.

This is a useful book g vin₀ practical guidance on writing oral tests suggesting test types, elicitation techniques, and marking systems and with one short chapter on basic concepts in test validity and reliability.

Valdman, Albert (ed.). 1988. *The Assessment of Foreign Language Oral Proficiency*. Special issue of *Studies in Second Language Acquisition*, Vol. 10, No. 2, June 1988.

This special issue contains articles on the ACTFL Scale and oral interview, the nature of proficiency, and other tests of oral proficiency.

Valdman, Albert. 1987. *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency*.

The book is in some nine sections dealing with the ACTFL Guidelines and their application to various languages and testing various aspects of language.

Valette, Rebecca M. 1967. *Modern Language Testing: A Handbook*. New York: Harcourt, Brace & World.

This book pre-dates developments in the concept and testing of proficiency and is included, like the books by Harris and Lado, as a standard reference in more traditional aspects of language testing. It discusses basic principles of test construction, administration and evaluation and has chapters on tests in the four macroskills and culture. An appendix lists with descriptions a number of commercially available tests in different languages (none in Japanese).

van Lier, Leo. 1989. "Reeling, Writhing, Drawling, Sketching and Fainting in Coils: Oral Proficiency Interviews as Conversation". In *TESOL Quarterly*, Vol. 23, No. 3, September 1989, pp. 489-508.

This paper provides a critique of the oral proficiency interview, looks at the differences between interviews and conversations, and discusses some of the major problems with interviews.

VanPatten, Bill. 1986. "The ACTFL Proficiency Guidelines: Implications for Grammatical Accuracy in the Classroom". In *Studies in Second Language Acquisition*, Vol. 8, No. 1, February 1986, pp. 56-67.

> One of the implications of proficiency and proficiency scales seems to be to play down the importance of grammatical accuracy in favour of more global assessment and communicative skills. This paper reviews the arguments for a focus on grammatical accuracy from the beginning of a language learning programme in the light of language acquisition research and the ACTFL Proficiency Guidelines and concludes that arguments in favour of such a focus are misguided.

Walker, Carolyn R. 1983. "The Whole is More than the Sum of its Parts: Evaluation Criteria in Oral Tests". In *British Journal of Language Teaching*, Vol. 21, No. 1, pp. 41-44.

> Walker considers the value of analytic as against global scales both from considerations of validity and reliability and from studies of the ARELS Oral Examination. She concludes that there is no advantage in the use of more analytic scales and their use puts unreasonable demands on examiners.

Young, Dolly J. 1986. "The Relationship between Anxiety and Foreign Language Oral Proficiency Ratings". In *Foreign Language Annals*, Vol. 19, No. 5, October 1986, pp. 439-445.

> This paper examines the effect of anxiety on ACTFL OPI results. It is shown to have a lesser effect than ability.

C. BIBLIOGRAPHY

[NB. This bibliography does not list items already listed in the Annotated Bibliography.]

Adams, Raymond J., Patrick Griffin, and Lyn Martin. 1987. "A Latent Trait Method for Measuring a Dimension in Second Language Proficiency". In *Language Testing*, Vol. 4, No. 1, 1987, pp. 9-27.

Alderson, J. Charles. 1987. *Innovation in Language Testing: Can the Microcomputer Help? Language Testing Update*, Special Report No. 1. Lancaster: University of Lancaster.

Alderson, J. Charles and Arthur Hughes. 1981. *Issues in Language Testing*, ELT Documents No. 111. London: British Council.

Alderson, J. Charles and A. H. Urquhart. 1985. "The Effect of Students' Academic Discipline on their Performance on ESP Reading Tests". In *Language Testing*, Vol. 2, No. 2, 1985, pp. 192-204.

American Council on the Teaching of Foreign Languages. 1987 *Applications of Oral Proficiency Testing to Foreign Language Instruction*. New York: ACTFL.

Antes, Richard L. 1985. "Teacher Education Admission Testing: Development and Implementation". (mimeograph) ED272498

Babcock, Arthur. 1985. "Proficiency Testing in French at the University of Southern California". In *The French Review*, Vol. 58, No. 3, February 1985, pp. 403-410.

Bachman, Lyle F. 1988. "Problems in Examining the Validity of the ACTFL Oral Proficiency Interview". In *Studies in Second Language Acquisition*, Vol 10, No. 2, June 1988, pp. 149-164.

Bachman, Lyle F. 1986. "The Test of English as a Foreign Language as a Measure of Communicative Competence". In Stansfield 1986.

Bachman, Lyle and Adrian S. Palmer. 1989. "The Construct Validation of Self-Ratings of Communicative Language Ability". In *Language Testing*, Vol. 6, No. 1, 1989, pp. 14-29.

Bachman, Lyle F. and Sandra J. Savignon. 1986. "The Evaluation of Communicative Language Proficiency: A Critique of the ACTFL Oral Interview". In *Modern Language Journal*, Vol. 70, No. 4, Winter 1986, pp. 380-390.

Baker, David. 1989. *Language Testing: A Critical Survey and Practical Guide*. London: Edward Arnold.

Barnwell, David. 1987. "Oral Proficiency Testing in the United States". In *British Journal of Language Teaching", Vol. 25, No. 1, Spring 1987, pp. 35-42.*

Briere, Eugene J. and Frances B. Hinofotis (eds.). 1979. *Concepts in Language Testing: Some Recent Studies*. Washington, D.C.: TESOL.

Brindley, Geoff. 1986. *The Assessment of Second Language Proficiency: Issues and Approaches*. Adelaide: National Curriculum Resource Centre.

Buck, Kathryn (ed.). 1989. *The ACTFL Oral Proficiency Interview Tester Training Manual*. Yonkers, NY: ACTFL.

Buckby, Michael et al. 1981. *Graded Objectives and Tests for Modern Languages: An Evaluation*. London: Schools Council.

Busch, D. 1982. "Introversion-Extroversion and the EFL Proficiency of Japanese Students". In *Language Learning*, Vol. 32, No. 1, 1982.

Byram, Michael. 1985. "The Relationship between Graded Objectives and Testing in Foreign Language Teaching and Language Awareness Work". In Donmall, B. G. (ed.). *Language Awareness*. York: NCLE; 1985.

Candlin, C. N. 1986. "Explaining Communicative Competence: Limits on Testability". In Stansfield 1986.

Carroll, Brendan J. 1982. "The Implications for Testing of the Communicative Approach to Language Teaching". In *Studies in Language Learning*, No. 4, pp. 45-55.

Carroll, Brendan J. and Patrick J. Hall. 1985. *Make your own Language Tests: A Practical Guide to Writing Language Performance Tests*. Oxford: Pergamon.

Clark, John L. D. 1986. "A Study of the Comparability of Speaking Proficiency Interview Ratings across Three Government Language Training Agencies". Washington, D.C.: Center for Applied Linguistics.

Clark, John L. D. 1983. "Language Testing: Past and Current Status - Directions for the Future". In *The Modern Language Journal*, Vol. 67, No. iv, 1983, pp. 431-441.

Clark, John L. D. 1982. "Some Sociolinguistic and Discourse Analysis Considerations in Oral Proficiency Interviewing". Paper to a pre-conference session on "Analysis of Spoken Discourse", Georgetown University Roundtable on Languages and Linguistics, Washington D.C.. ED247762 FL014530

Clark, John L. D. 1972. *Foreign Language Testing: Theory to Practice*. Philadelphia, Pa.: Center for Curriculum Development.

Clark, John L. D. and Li Ying-che. 1986. *Development, Validation, and Dissemination of a Proficiency-based Test of Speaking Ability in Chinese and an Associated Assessment Model for Other Less Commonly Taught Languages*. Washington, D.C.: Center for Applied Linguistics.

Cohen, Andrew D. 1980. *Testing Language Ability in the Classroom* Rowley, Mass.: Newbury House.

Cohen, Andrew D. 1984. "On Taking Language Tests: What the Students Report". In *Language Testing*, Vol. 1, NO. 1, 1984, pp. 70-81.

Cziko, Gary. 1984. "Some Problems with Empirically-based Models of Communicative Competence". In *Applied Linguistics*, Vol. 5, No. 1, 1984, pp. 23-37.

Cziko, Gary A. 1982. "Improving the Psychometric, Criterion-Referenced, and Practical Qualities of Integrative Language Tests". In *TESOL Quarterly*, Vol. 16, No. 3, September 1982, pp. 367-379.

Davies, Alan (ed.). 1968. *Language Testing Symposium: A Psycholinguistic Approach*. London: Oxford.

Delamere, T. 1985. "Notional-Functional Syllabi and Criterion-Referenced Tests: The Missing Link". In *System*, Vol. 13, 1985.

Douglas, Dan and Larry Selinker. 1985. "Principles for Language Tests within the 'Discourse Domains' Theory of Interlanguage: Research, Test Construction and Interpretation". In *Language Testing*, Vol. 2, No. 2, pp. 205-226.

Educational Testing Service. 1986. *Oral Proficiency Familiarization Kit*. Boston, Ma.: Heinle & Heinle.

Evans, C., J. Gaudin, and F. Raveau. 1987. "Testing Advanced Communicative Competence through Role-Play". In *British Journal of Language Teaching*, Vol. 25, No. 2, 1987.

Foreign Service Institute. 1986. "Oral Language Proficiency Testing at the Foreign Service Institute". Washington, D.C.: FSI School of Language Studies.

Green, P.S. and K. Hecht. 1987. "The Influence of Accuracy on Communicative Effectiveness". In *British Journal of Languge Teaching*, Vol. 25, No. 2, 1987.

Hale, G. A., Charles W. Stansfield, and R. P. Dwan. 1984. "A Comprehensive TOEFL Bibliography, 1963-82". In *The Modern Language Journal*, Vol. 68, No. 1, 1984.

Harrison, Andrew. 1983. *A Language Testing Handbook*. London: Macmillan.

Henning, Grant. 1987. *A Guide to Language Testing: Development, Evaluation, Research*. Rowley, Mass.: Newbury House.

Henning, Grant, Thom Hudson and Jean Turner. 1985. "Item Response Theory and the Assumption of Unidimensionality for Language Tests". In *Language Testing*, Vol 2, No. 2, 1985, pp. 141-154.

Hijirida, Kyoko. 1989. "ACTFL no nerai: Jituyo noryoku o ika ni sodateru ka" (Goals of ACTFL: How to Cultivate Practical Abilities). In *Nihongo*, June 1989.

Hirsch, Bette G. 1985. "A Proficiency-Based French Conversation Course". In *The French Review*, Vol. LIX, No. 2, December 1985, pp. 210 - 218.

Hughes, Arthur and Don Porter (eds.). 1983. *Current Developments in Language Testing*. London: Academic Press.

Hyltenstam, Kenneth and Manfred Pienemann. 1985. *Modelling and Assessing Second Language Acquisition*. Clevedon: Multilingual Matters.

Ingram, D. E. 1990. "The International English Language Testing System (IELTS): Its Nature and Development". Paper to the 1990 RELC Seminar, Regional Language Centre, Singapore, 9-12 April, 1990. (mimeograph)

Ingram, D. E. 1990a. "The International English Language Testing System (IELTS): The Speaking Test". Paper to the 1990 RELC Seminar, Regional Language Centre, Singapore, 9-12 April, 1990. (mimeograph)

Ingram, D. E. 1985. "Assessing Proficiency: An Overview on Some Aspects of Testing". In Hyltenstam and Pienemann 1985: 215-276.

Ingram, D. E. 1984. "Using Proficiency Rating Scales with High School Foreign Language Learners". Paper to the Biennial National Languages Conference of AFMLTA, Hobart, 31 August to 4 September, 1984. ED249767 FL014540

Ingram, D. E. and Elaine Wylie. 1979/1985. *The Australian Second Language Proficiency Ratings*. (mimeograph). Also Canberra: DIEA/AGPS; 1984.

Ingram, D. E. 1979/1982. "An Introduction to the Australian Second Language Proficiency Ratings". (mimeograph). Also Canberra: DIEA/AGPS; 1984.

Jarvis, G. C. 1986. "Proficiency Testing: A Matter of False Hopes?" In *ADFL Bulletin*, Vol. 18, No. 1, 1986.

Johnson, K. and D. Porter. 1983. *Perspectives in Communicative Language Teaching*. London: Academic Press.

de Jong, John H.A.L. and C. A. W. Glas. 1987. "Validation of Listening Comprehension Tests using Item Response Theory". In *Language Testing*, Vol. 4, No. 2, December 1987, pp.170-194.

Kaplan, Isabelle M. 1984. "Oral Proficiency Testing and the Language Curriculum: Two Experiments in Curricular Design for Conversation Courses". In *Foreign Language Annals*, Vol. 17, No. 5, 1984, pp. 491-497.

Keitges, D. J. 1982. "Language Proficiency Interview Testing: An Overview". In *JALT Journal*, Vol. 4, 1982, pp. 17-45.

Kramsch, Claire J. 1986. "Proficiency vs. Achievement: Reflections on the Proficiency Movement". In *ADFL Bulletin*, Vol. 18, No. 1, 1986, pp. 22-24.

Krzanowski, Wojtek J. and Anthony J. Woods. 1984. "Statistical Aspects of Reliability in Language Testing". In *Language Testing*, Vol. 1, No. 1, 1984, pp. 3-19.

Lantolf, James P. and Frawley, William. 1988. "Proficiency: Understanding the Construct". In *Studies in Second Language Acquisition*, Vol. 10, No. 2, 1988, pp. 181-195.

Low, Graham D. 1982. "The Direct Testing of Academic Writing in a Second Language". In *System*, Vol. 10, No. 3, 1982, pp. 247-257.

Lowe, Pardee. 1983. "The ILR Oral Interview: Origins, Applications, Pitfalls and Implications". In *Die Unterrishtspraxis*, Vol. 16, No. 2, 1983.

Lowe, Pardee and J. E. Liskin-Gasparro. 1982. *Testing Speaking Proficiency: The Oral Interview*. Washington, D.C.: Center for Applied Linguistics.

Lowe, Pardee and Charles Stansfield (eds.). 1988. *Second Language Proficiency Assessment: Current Issues*. Englewood Cliffs, NJ: Prentice Hall Regents.

Lange, Dale L. and Pardee Lowe. 1988. "Rating Reading Passages according to the ACTFL Reading Proficiency Standard: Can it be Learned?". In *Foreign Language Annals*, Vol. 21, No. 3, May 1988, pp. 227-239.

Madsen, Harold S. 1983. *Techniques in Testing*. Oxford: University Press.

Magnan, Sally S. 1987. "Rater Reliability of the ACTFL Oral Proficiency Interview". In *Canadian Modern Language Review*, Vol. 43, No. 3, March 1987, pp. 525-537.

Makino, Seeichi. 1987. "ACTFL gengo noryoku kijun to Amerika ni okeru nihongo kyoikuu" (ACTFL Proficiency Guidelines and Japanese Language Education in the U.S.). In *Nihongo Kyoiku*, No. 61, March, 1987.

Massad, C. E. 1980. *Language Proficiency Tests Developed for the IEA International Study of Achievement in English as a Foreign Language, ELT Document No. 108*. London: British Council.

McBride, John. 1990. "Computer Assisted Instruction in the Teaching of Japanese as a Second Language: An Issues Paper prepared for an Asialink Search Seminar held at the University of Wollongong, 24 May, 1990". (mimeograph)

Morrow, Keith. 1990. "Evaluating Communicative Tests". Paper to the 1990 RELC Seminar, Regional Language Centre, Singapore, 9-12 April, 1990. (mimeograph)

Morrow, Keith. 1985. "The Evaluation of Tests of Communicative Performance". In *Prospect*, Vol. 1, No. 2, 1985, pp. 32-44.

Nunan, David. 1987. "The Ghost in the Machine: An Examination of the Concept of Proficiency". In *Prospect*, Vol. 2, No. 2, 1987.

Oller, John W. 1979. *Language Tests at School*. London: Longman.

Oller, John W. 1983. *Issues in Language Testing Research*. Rowley, Mass.: Newbury House.

Oller, John W. 1974. "Expectancy for Successive Elements: Key Ingredient to Language Use:. In Foreign Language Annals, Vol. 7, 1974: 443-452.

Oller, John W. 1973. "Cloze Tests of Second Language Proficiency and What they Measure". In Language Learning, Vol. 23, No. 1, June 1973: 105-118.

Oller, John W. 1973a. "Discrete Point Tests versus Tests of Integrative Skills". In Oller and Richards 1973: 184-199.

Oller, John W. and Jack C. Richards. 1973. Focus on the Learner: Pragmatic Perspectives for the Language Teacher. Rowley, Mass.: Newbury House.

Omaggio, Alice C. 1985. *Proficiency, Curriculum, Articulation: The Ties that Bind*. Middlebury, VT.: North East Conference on the Teaching of Foreign Languages.

Omaggio, Alice C. 1983. *Proficiency-Oriented Classroom Testing, Language in Education: Theory and Practice, No. 52*. Washington, D.C.: Center for Applied Linguistics.

Palmer, Adrian S. 1981. "Testing Communication". In *IRAL*, Vol 10, 1981, pp. 35-45.

Palmer, Adrian S. *The Construct Validation of Tests of Communicative Competence*. Washington, D.C.: TESOL.

Palmer, Adrian S., Peter J. M. Groot, and George A. Trosper. 1981. *The Construct Validation of Tests of Communicative Competence*. Washington, D.C.: TESOL.

Pienemann, Manfred and M. Johnston. 1986. "An Acquisition-Based Procedure for Second Language Assessment". In *Australian Review of Applied Linguistics*, Vol. 9, No. 1, 1986.

Pienemann, Manfred, Malcolm Johnston and Geoff Brindley. 1988. "Constructing an Acquisition-based Procedure for Second Language Assessment". In *Studies in Second Language Acquisition*, Vol. 10, 1988, pp. 217-243.

Piper, Terry H. and P. F. Leicester. 1980. "Word Association as an Indicator of English Language Proficiency". (mimeograph) ED227651

Politzer, R. L. and M. McGroarty. 1983. "A Discrete Point Test of Communicative Competence". In *IRAL*, Vol 21, No. 3, 1983, pp. 179-191.

Regents Examinations, Regents Competency Tests, and Proficiency Examinations: School Administrator's Manual. New York: University of the State of New York, State Education Department and Bureau of Elementary and Secondary Testing Programs; 1989.

Reisener, H. 1983. "Testing Communicative Ability". In *Practical English Teaching*, Vol. 3, No. 3, 1983.

Reschke, C. 1984. "Oral Proficiency Testing: Meeting a Variety of Needs". In *Die Unterrichtspraxis*, Vol. 17, 1984, pp. 265-278.

Rosser, Harry. 1986. "Language Proficiency Assesment: Implications and Applications". In *ADFL Bulletin*, Vol. 18, No. 1, 1986, pp. 15-19.

Schulz, Renate A. 1988. "Proficiency-Based Foreign Language Requirements: A Plan for Action". In *ADFL Bulletin*, Vol. 19, No. 2, January 1988, pp. 24-28.

Shohamy, Elana. 1983. " The Stability of Oral Proficiency Assessment on the Oral Interview Testing Procedures". In *Language Learning*, Vol. 33, No. 4, 1983, pp. 527-540.

Skehan, Peter. 1983. "Issues in the Testing of ESP". In *MALS Journal*, Vol. 8, pp. 61-85.

Skehan, Peter. 1984. "Issues in the Testing of English for Specific Purposes". In *Language Testing*, Vol. 1, No. 2, pp. 202-220.

Spolsky, Bernard. 1989. "Communicative Competence, Language Proficiency, and Beyond". In *Applied Linguistics*, Vol 10, No. 2, pp. 138-156.

Spolsky, Bernard. 1985. "What does it Mean to Know How to Use a Language? An Essay on the Theoretical Basis of Language Testing". In *Language Testing*, Vol. 2, No. 2, pp. 180-191.

Stansfield, Charles W. 1989. *Simulated Oral Proficiency Interviews. ERIC Digest*, December 1989. Washington, D.C.: Center for Applied Linguistics.

Stansfield, Charles W. (ed.). 1986. *Toward Communicative Competence Testing: Proceedings of the Second TOEFL Invitational Confernce, TOEFL Research Reports No. 21*. Princeton, NJ: Educational Testing Service.

Stansfield, Charles W. 1986a. *Technology and Language Testing*. Washington, D.C.: TESOL.

Stansfield, Charles W. and Dorry Mann Kenyon. 1989. "A Report on Extending the ACTFL Proficiency Guidelines to the Less Commonly Taught Languages". In *Newsletter of the NorthEast Conference on the Teaching of Foreign Languages*, No. 26, pp. 28-30.

Stevenson, Kiyo Yamada. 1989. Opi shikenkan yosei no jittai repoto" (A report on OPI Tester Training). In *Nihongo*, June 1989.

Sunderland, Jane et al. 1988. "To Be an Explainer: A Reporting Test of Oral Proficiency". In *ELT Journal*, Vol. 42, No. 1, January 1988, pp. 28-33.

Takashima, Hideyuki. 1987. "To What Extent are Non-Native Speakers Qualified to Correct Free Composition? A Case Study". In *British Journal of Language Teaching*, Vol. 25, No. 1, 1987, pp. 43-48.

Thompson, Irene, Richard T. Thompson and David Hiple. 1987. "Adapting the ACTFL/ETS Guidelines to the Less Commonly Taught Languages". In Stansfield and Harman 1987, Chapter 4.

Thompson, Richard T. and Dora E. Johnson. 1988. *ERIC Digest: Proficiency Testing in the Less Commonly Taught Languages*. Washington, D.C.: Center for Applied Linguistics.

Usami, Mayumi. 1989. "OPI shikenkan yosei koza juko repoto" (A Trainee's Report on an OPI Tester Training Workshop). In *Nihongo*, June 1989.

Vollmer, Helmut J. 1981. "Why are we interested in 'General Language Proficiency'?". In Alderson and Hughes 1981: 152-175.

Wesche, Marjorie Bingham. 1983. "Communicative Testing in a Second Language". In *The Modern Language Journal*, Vol 67, No. 1, 1983, pp. 41-55.

D. RELEVANT TESTS, SCALES etc.

American Council of Teachers of Foreign Languages. 1982/1986. *ACTFL Proficiency Guidelines*. New York: ACTFL.

These scales were derived from the earlier published scales with clear influences from FSI, ILR and ASLPR. They provide a "hierarchy of global characterizations of integrated performance" in the four macroskills individually. They are intended to measure global performance using short descriptors for up to ten different levels in each macroskill. Versions are available for, amongst others, Chinese, Russian and Japanese as well as the generic English version.

Buros, Oscar K. (ed.). 1975. *Foreign Language Tests and Reviews*. Highland Park, NJ.: Gryphon Press.

Though the most recent edition seems to be 1975, this remains a useful listing of tests and reviews of them in areas relevant to language learning and teaching.

Educational Testing Service. 1981/1986. *Japanese Proficiency Test*. Princeton, N.J.: Educational Testing Service.

This norm-referenced test is used by many schools in the United States. It is intended for use with native English speakers learning Japanese as a foreign language with proficiencies ranging from that achieved after one year's study to that "associated with a college or graduate-level language program". It provides scores in listening and reading comprehension together with a diagnostic score for Japanese character recognition. The test is independent of any course and uses realistic language-use situations and contemporary texts. It uses fairly standard multiple-choice type questions to elicit comprehension levels. The results are a score between 20 and 80. A score-comparison table provides the distribution of scores for each of the three areas (listening comprehension, reading comprehension and character recognition) though no indication of the nature of the population (other than by "beginner", "intermediate" and "advanced") is provided. Similarly, there is no attempt made to provide a descriptive scale to facilitate interpretation of results.

The Japan Foundation. 1984/1989. *The Japanese-Language Proficiency Test*. Tokyo: Japan Foundation/Association of International Education, Japan.

This norm-referenced test was developed by the Japan Foundation and the Association of International Education, Japan as the basis on which the proficiency of foreign students of Japanese may be certified. It is available once per year simultaneously in Japan and abroad. It is primarily intended for students wishing to go to Japan to study. The test has four levels ranging (according to the brief descriptive criteria provided) from Level 4 (which requires a

knowledge of 100 words of kanji, 800 words of vocabulary, and the ability to engage in simple conversations) to Level 1 (where 2000 kanji and 10,000 words of vocabulary are expected together with the ability to survive in academic contexts). Each level has three sections: "writing-vocabulary", "hearing" and reading-grammar". "Passing scores" are set at 60% or 70% though no information is provided to indicate what this means in performance terms though Level 1 is necessary for entry to a Japanese university. The test is designed to be machine marked and is multiple-choice in format. The test seems to require a high level of "test-taking" expertise with long and quite complicated instructions as to how to fill in the answer sheets. However, it does not meet any of the criteria we have come to expect for a genuine proficiency test. In particular, the item format is not conducive to realistic language use and hence to the assessment of practical proficiency, it cannot be regarded as assessing productive skills, the results are, at best, difficult to interpret in terms of language behaviour, performance or practical proficiency and, as with any "indirect" test, one cannot readily identify the criteria on which such decisions are made as whether or not an applicant's proficiency is adequate for academic study in Japan. The information that accompanies the test provides extensive data on how many candidates from what countries took the test and their distribution of scores. Though this test has acquired considerable status, one cannot regard it as an adequate proficiency test and one would have to expect that it has the same shortcomings when used as such as do other "indirect" tests. (See also the entry under Munson and Harrison in the Annotated Bibliography above.)

Waites, Carol and Rena Kelly. 1989. "Towards TOLSKIP: A Review of Selected Language Testing Systems in relation to the Needs of the Australian Tourism Industry - Report to the Australia-Japan Foundation". (mimeograph).

This report reviews a number of instruments in relation to their usefulness in meeting the needs of the tourism industry for measures of language proficiency. Though some testers would wish to query certain of the statements made for their accuracy of fact or interpretation, it is a useful list of instruments developed in Australia in recent years.

PROGRAM

DAY 1 - TUESDAY JUNE 12

8.30 am - 9.00 am	Registration and administration.
9.00 am - 9.05 am	Welcome and introduction of Dr Stephen FitzGerald, Chairman, Asian Studies Council.
9.05 am - 9.20 am	Opening address by Dr Stephen FitzGerald.
9.20 am - 10.00 am	Presentation by Professor David Ingram - Language Proficiency Testing in the International Setting. (Overview statement setting goals for International Working Group and addressing points in the background overview paper distributed to participants before the meeting).
10.00 am - 11.00 am	The Employers' Perspective. Panel of representatives of industry.
11.00 am - 11.15 am	Morning tea.
11.15 am - 12.45 pm	Plenary session: Response to speakers.
12.45 pm - 1.30 pm	Lunch.
1.30 pm - 3.00 pm	International Overview. Presentations by Mr Masaya Usuda, Professor Kazuo Otsubo and Professor Kunio Okuda (Japan), Mr Tom Whalley (Canada) and Mr David Hiple and Dr Hiroko Kataoka (U.S.A.). (Presentations providing a brief overview of proficiency assessment and developments in Japanese language proficiency testing in the speaker's country).
3.00 pm - 3.15 pm	Afternoon tea.
3.15 pm - 3.50 pm	Presentation by Professor Jiri Neustupny - Language Planning and the Testing of Japanese in Australia.
3.50 pm - 4.30 pm	Plenary session to suggest future trends for the assessment of proficiency in the Japanese language in Australia within the context of international trends.

4.30 pm - 5.15 pm	Workshop discussions to develop a statement on overall needs for Japanese language proficiency assessment in Australia.
5.15 pm - 6.00 pm	Small working group to plan tasks for Day 2 arising out of the conclusions reached on Day 1.

DAY 2 - WEDNESDAY JUNE 13

9.00 am - 9.30 am	Presentation by Professor Alan Davies - Designing Instruments to Measure Language Proficiency.
9.30 am - 10.20 am	Report back to plenary by small working group.
10.20 am - 10.35 am	Morning tea.
10.35 am - 12.05 am	Workshops to identify assessment instruments to be developed in response to needs identified and to sketch initial specifications.
12.05 am - 1.05 pm	Report back to plenary by workshop groups on discussions and conclusions.
1.05 pm - 1.50 pm	Lunch.
1.50 pm - 2.20 pm	Plenary session to resolve assessment instruments to be developed.
2.20 pm - 3.50 pm	Workshops to develop specifications for agreed assessment instruments.
3.50 pm - 4.05 pm	Afternoon tea.
4.05 pm - 5.30 pm	Report back to plenary by workshop groups on discussions concerning specifications. Plenary session to resolve specifications.

DAY 3 - THURSDAY JUNE 14

9.00 am - 9.30 am	Short plenary session to outline tasks for the day.
9.30 am - 10.30 am	Workshops to develop details of projects.
10.30 am - 10.45 am	Morning tea.
10.45 am - 12.45 pm	Workshops continuing.
12.45 pm - 1.30 pm	Lunch.
1.30 pm - 5.00 pm	Working groups to summarise outcomes and finalise proposals.

PARTICIPANTS

Mrs Sherrill Bell
Executive Director
NAATI
P O Box 349
JAMISON CENTRE ACT 2614

Mr Geoff Brindley
National Centre for English Language
Macquarie University
NORTH RYDE NSW 2113

Ms Ajoika Chiaki
Japanese Sub-titler
SBS
P O Box 028
MILSONS POINT NSW 2061

Mr Peter Davidson
Director
Japanese Proficiency Unit
Queensland University
ST LUCIA QLD 4067

Professor Alan Davies
Director
Language Testing Unit
Department of Linguistics
and Language Studies
University of Melbourne
PARKVILLE VIC 3052

Dr Patrick Griffin
Phillip Institute of Technology
Alva Grove
COBURG VIC 3058

Mr David Hiple
Project Director
American Council of Teaching &
Foreign Languages, Inc. (ACTFL)
P O Box 1077
YONKERS NY 10701-6801

Mr Fumio Isoda
First Secretary
Japanese Embassy
112 Empire Circuit
YARRALUMLA ACT 2601

Dr Hiroko Kataoka
Associate Professor of Japanese
Department of East Asian Languages
and Literature
University of Oregon
EUGENE

Ms Rena Kelly
Director
Institute of Languages
University of New South Wales
P O Box 1
KENSINGTON NSW 2033

Mr R T Kinnaird
R T Kinnaird and
Associates Pty Ltd
4/25 Marshall Street
MANLY NSW 2095

Mr Tony Kitchen
Queensland Catholic Education
Commission
143 Edward Street
BRISBANE QLD 4000

Ms Jenny Lambert
Industrial Research Officer
Australian Hotel Association
Level 5
9 Quay Street
SYDNEY NSW 2000

Dr John McArthur
Secretary
Australian Education Council
GPO Box 349
MELBOURNE VIC 3001

Ms Roslyn McLeod
Managing Director
Tour Hosts
GPO Box 128
SYDNEY NSW 2001

Ms Inara Merrick
Department of School Education
GPO Box 33
SYDNEY NSW 2001

Ms Pam Moss
Curriculum Policy Branch
WA Ministry of Education
151 Royal Street
EAST PERTH WA 6000

Professor Jiri Neustupny
Professor of Japanese Studies
Japanese Studies Centre and
Institute for Contemporary
Asian Studies
Faculty of Arts
Monash University
CLAYTON VIC 3168

Professor Kazuo Otsubo
5/546-2 Matsushiro, Tsukuba-shi
Ibaraki-ken
JAPAN 305
c/- Mr Masuharu Yamada
Head
Japanese Studies Department
Japan Foundation
Park Building
306, Kioi-choi
Chiyoda-ku
TOKYO 102 JAPAN

Professor Kunio Okuda
41-9-8, 4-Chome, Yasuhigashi
Asaminami-ku
Hiroshima-shi
JAPAN
c/- Mr Masaharu Yamado
Head Japanese Studies Department
Japan Foundation
Park Building
3-6, Kioi-choi
Chiyoda-ku
TOKYO 102 JAPAN

Ms Sharon O'Donnell
Human Resource Development
Section
Qantas Airways Limited
GPO Box 489
SYDNEY NSW 2000

Mr Marcel Poualion
Acting Director
Division of Operational Services
Bureau of Employment, Vocational
and Further Education and Training
18th Floor State Law Building
50 Ann Street
BRISBANE QLD 4000

Mr John A Pitman
Chairman
Australian Committee of
Chief Executive Officers
Australasian Conference
of Assessment and
Certification Authorities
Board of Senior Secondary
School Studies
P O Box 307
SPRING HILL QLD 4004

Mr Noel Salmon
Manager
Language Training Unit
Department of Foreign
Affairs and Trade
Administrative Building
Parkes Place ACT 2600

Professor Masami Sekine
Japanese Embassy
112 Empire Circuit
YARRALUMLA ACT 2601

Ms Robyn Spence-Brown
Japanese Studies Centre and
Institute for Contemporary Asian
Studies
Faculty of Arts
Monash University
CLAYTON ACT 3168

Dr Paul Tuffin
NAFLaSSL Senior Project Officer
c/- SABSA
60 Greenhill Road
WAYVILLE SA 5034

Mr Masaya Usuda
c/- Mr Masaharu Yamada
The Japan Foundation
Park Building
3-6 Kioi-choi
Chiyoda-ku
TOKYO 102 JAPAN

Mr Laurie Westcott
Department of Immigration,
Local Government and
Ethnic Affairs
Benjamin Offices
Chan Street
BELCONNEN ACT 2616

Mr Bill Wise
Director
The Australia-Japan Foundation
PO Box E500
Queen Victoria Terrace
CANBERRA ACT 2600

Planning Committee

Mr Martyn Cove
Deputy Chairman
Asian Studies Council
c/- Tas Department of Education
116 Bathurst Street
HOBART TAS 7000

Ms Elaine Wylie
Institute of Applied Linguistics
Qld University of Technology
Kelvin Grove Campus
Victoria Park Road
Kelvin Grove
BRISBANE QLD 4059

Mr Joe Lo Bianco
Director
Languages Institute of
Australia
c/- Monash City Offices
41-43 Exhibition Street
MELBOURNE VIC 3000

Ms Jenny Shapcott
Program Manager and
Japan Advisor
Asian Studies Council

ASC Chairman
Dr Stephen FitzGerald
ASC Chairman

Ms Carol Waites
Head of Modern Languages
Institute of Languages
University of New South Wales
P O Box 1
KENSINGTON NSW 2033

Mr Tom Whalley
Coordinator, Pacific Rim Languages
Modern Languages and Multicultural
Programs
British Columbia Ministry of
Education
Fax: 604.387-1470

Interpreters
Ms Mari Kitasaka

Ms Emiko Yamaguchi
c/- Conference Interpreter Services
GPO Box 128
SYDNEY NSW 2001

Professor David Ingram
Centre for Applied Linguistics
and Languages
Division of Asian and International
Studies
Griffith University
NATHAN QLD 4111

Mr Jim Wilson
National Curriculum Advisor
Asian Studies Council
c/- SA Department of Education
5th Floor, Education Centre
31 Flinders Street
ADELAIDE SA 5000

Mr Tim McNamara
Department of Linguistics and
Language Studies
Language Testing Unit
University of Melbourne
PARKVILLE VIC

Mr Roger Peacock
Executive Director
Asian Studies Council

ASC Secretariat Support
Project Officer
Asian Studies Council Secretariat

OVERVIEW PAPER: TOWARDS THE DEVELOPMENT OF PROFICIENCY AND OTHER TESTS IN JAPANESE AS A FOREIGN LANGUAGE IN AUSTRALIA

D. E. INGRAM

Griffith University, Brisbane

CONTENTS

I Preamble

II The Nature of Proficiency and its Assessment

III Assessment of Proficiency

IV Needs that Japanese Tests and Scales Must Meet in Australia

- Battery
- Proficiency
- Industry
- Language Programmes
- Quality Control
- Diversifying and Increasing Language Learning
- Attainment Tests
- Streaming
- Diagnosis
- Self-Assessment
- Classifying Materials and Activities
- Programme Rationalization
- Other Tests
- Research
- Training in Testing

V Test Types

VI Tests of Japanese

VII The Future

[N.B. The references for this paper are contained in one of the bibliographies prepared for this conference and reprinted elsewhere in this volume.]

I **PREAMBLE**: In preparation for this conference, I was asked to provide an overview of language proficiency assessment in recent years and, specifically, to provide an annotated bibliography of significant works, an unannotated bibliography, and a background paper which would review the literature. I must say that I approached the task with some trepidation since, through the 1980's, the literature on proficiency and its assessment has proliferated. Indeed few concepts in applied linguistics have generated so much scepticism when first explicitly introduced and subsequently so much interest and activity. The scepticism, which is not entirely in the past, is difficult to explain unless one recognizes the seductive hold that the false objectivity and false security of statistics and psychometrics have exerted over language testing. The reasons for the interest and the activity are more readily explained since proficiency is at the very heart of what language teaching is about: whatever else language teachers aspire to do, they must seek the development of proficiency, that is, they must seek to develop in learners the ability to use the language for practical communication tasks. In fact, virtually everything else that language teachers have both traditionally and more recently sought in language teaching (literary appreciation, aesthetic response, cultural understanding, translation skills, ability to use the language in everyday or in specialized situations, and so on), all are dependent on the learners' developing the ability to use the language. To be fair to the sceptics, it has also to be said that some of their scepticism arises from academics' legitimate fear of things that cannot be unambiguously defined: though everyone, both applied linguists and laypersons, uses the term "proficiency" and knows what it means when they say that "Bill is proficient in French" or "Mary is proficient in Japanese", it is in fact difficult to arrive at a fully explicit definition of the concept and this seems to make the development of tests to measure it more difficult and more open to attack.

The other reason why I approached with trepidation the task of preparing a review paper for this conference was that, having prepared quite lengthy bibliographies, I felt that the review paper might become somewhat repetitious. In addition, I had already some years ago written a lengthy paper looking at the different types of tests and assessing their usefulness for the measurement of proficiency and, more recently, Elaine Wylie and I had written a paper in which we reviewed the different approaches that could be adopted in developing proficiency scales: both of these papers are in the portfolio. Finally, I was also conscious of the fact that the audience to whom I would be addressing this paper was highly select, being either laypersons who are end-users of language skills and test results or language testers who could be expected to be very familiar with the concept of proficiency, its assessment and the associated literature.

Consequently, in preparing for this conference, I have sought to survey the literature as comprehensively as possible, select a cross-section of items for inclusion in the annotated and unannotated bibliographies, and survey whatever tests of Japanese proficiency are already available; I have tried to provide in the portfolio, papers which were immediately relevant to the task of test development, which would exemplify some of the different proficiency instruments available and raise some key issues to be considered by persons charged with developing tests. Consequently, this paper does not attempt to review the literature (the bibliographies do that) but it aims to provide an overview of the issues that are relevant to the purpose of this meeting and of the working party, to point to the directions in which our thoughts and discussions might run, and to suggest

the broad issues to be considered as we seek to devise ways by which the field of language testing and of proficiency assessment in particular might respond to the need for Japanese language skills in Australia both now and into the 21st Century.

II THE NATURE OF PROFICIENCY AND ITS ASSESSMENT: In the paper on "Assessing Proficiency" in the portfolio, it was observed that, over the last forty years, the concept of proficiency and the approach to measuring it have changed as our understanding of the nature of language and of language learning has changed. When language was seen as a set of rules and vocabulary that were memorized and consciously applied, proficiency in a language was equated with knowing the grammatical rules and lists of vocabulary and being able to formally apply them. From here, "proficiency" passed through behaviourist notions of an accretion of conditioned habits, then the internalized rules of cognitivism, until, as functional-notional and communicative ideas developed, applied linguists and testers departed from the analytic approach, realized that language is more than the sum of its parts, and moved the focus onto the tasks that learners can carry out in the language and how they are carried out. In other words, having passed through the "traditional", the behaviourist, the cognitivist, the functional-notional and the communicative stages, the definition of proficiency now is seen to embrace both the learners' ability to carry out tasks in the language and how those tasks are carried out; rather than being analytic, the approach is now synthetic and global, with the focus on the learners' total language behaviour or performance. Characteristically this behaviour is described in scales which seek to capture the sort of language behaviour or performance that can be observed in learners as their language develops from zero through to native-like. This also means that, whereas more traditional approaches to testing are analytic in that they seek to analyse the language into its component parts and measure proficiency by adding up the learners' scores on each of the discrete bits of the language, the overriding though not exclusive emphasis today in assessing proficiency is on the learners' overall or global performance with their knowledge of individual grammatical rules, vocabulary items or other elements of the language primarily of interest insofar as they contribute to overall language behaviour.

There have been many attempts to define proficiency and to capture its principal features. Some of these are referred to in the paper "Assessing Proficiency". Vollmer [1981] seems to opt out and defines proficiency as what proficiency tests measure. Politzer and McGroarty [1983] define it simplistically as the ability to convey or receive information, which seems to omit consideration of the way in which the information is conveyed but, in fact, they claim to show in their study that, though communicative competence is distinct from linguistic competence and the learner may have strategies that enable him or her to compensate for deficiencies in linguistic competence, linguistic competence nevertheless imposes limits on communicative competence. Perhaps the most cited characterisation of proficiency is that of Canale and Swain [1980] in which they identify four components, viz., grammatical competence (in the widest sense of "grammar" to entail all the rules by which utterances are put together), sociolinguistic competence (or knowledge of the sociocultural rules of use and of discourse), and strategic competence (or the set of verbal and non-verbal communication strategies that learners adopt "to compensate for breakdowns in communication due to performance variables or to insufficient competence") [Canale and Swain 1980: 30; Swain 1984: 9]. Bachman [1988] defines proficiency as the ability to conduct culturally-embedded verbal interaction with speakers of the target language and he identifies a

similar array of underlying features as do Canale and Swain with, in addition, reference to motor skills. An important distinction made by Griffin and others [1988] and made also in the literature on the ASLPR [e.g. Ingram 1979/1982, 1984 in unannotated bibliography, 1985] is that between course-dependent performance using the material taught during the course and general proficiency or, in the terms used by Griffin et al, "the effective use of language in contexts...irrespective of the body of material taught in class" [1988:1]. A further essential distinction is made in "Assessing Proficiency" between proficiency and "communicative competence" in the loose sense of the ability to communicate: the ability to communicate and the success of attempts at communication depend on things beyond facility in the language itself including such factors as intelligence, knowledge and the ability to organize it, introversion or extroversion, and the willingness of the interlocutor to accommodate less than native-like language. Proficiency (at least in the sense in which it is used in the documents on the ASLPR [Ingram 1979/82, 1985]) relates to facility in the language as it is used to carry out language tasks.

III ASSESSMENT OF PROFICIENCY: Approaches to testing or the assessment of proficiency have also differed according to the view of language and its learning dominant at the time. The analytic approaches referred to above, described in "Assessing Proficiency" (in Sections II.2 and III.2), and characteristic of so-called indirect tests assessed proficiency by analysing the language into its component parts and assessing the learners' knowledge of or ability to use those component parts. Although such approaches are very effective for diagnostic testing, in assessing proficiency the problem is, first, that language is more than the sum of its component parts and, second, the sum of the items scored correctly on a test does not necessarily give much indication, or any readily interpretable indication, of how well the learners can use the language in non-test, real-life situations.

Predominantly in the late-1960's and the 1970's, there was much interest (illustrated especially in Oller's work [e.g. 1979, 1983]) in developing test types that focus on discrete items in the language but do so in a total language context. These tests, which drew on the redundancy of language and, in Oller's terms [1973, 1973a, 1974], on the learner's expectancy grammar, sought to be as objective as indirect tests, to focus on individual items, but to embed them in a full text. Characteristic of these tests are dictation, white noise, interlinear and the various types of cloze but, as with indirect tests, their interpretation is difficult and, however effective they might be for some diagnostic purposes and for rank-ordering learners, it is difficult to move from the scores to an unambiguous statement about what the learners can do in the language.

To overcome problems of interpretation and to try to assess proficiency in situations where the language is operating more naturally (i.e. with all its elements contributing to carrying out language tasks, to achieving some purpose, and in a sociocultural context), approaches have been sought that allow the language to operate naturally in a communication situation in which the learner is carrying out normal language tasks. The assessment process is a matter of rating the observed language behaviour against a scale which describes learners' language behaviour as it develops from zero to native-like. This is commonly done by observing the learners' use of the language in an interview designed for the purpose, during which the observed behaviour is matched against the scale descriptions. However, other approaches are also possible. Some testers [e.g. Brindley 1986, Pieneman et al 1988] see a close relationship between proficiency levels

and the emergence of certain elements of the language (especially syntax) and devise tests that focus on these features in order to relate learners to proficiency levels. Rating using proficiency scales is essentiallly a process of observing language behaviour and matching it against scale descriptions and this observation or, more accurately, the process of eliciting language behaviour to observe, may be encapsulated in more traditional types of tests with either the scores or performance on individual items related in some way to the scale levels. Ellis [1984] seems to abandon the attempt to identify the nature of proficiency, theorizes that the greater the learners' proficiency the fewer "communication strategies" they will use, and attempts to use this phenomenon to rate learners' proficiency. Others [e.g. Politzer and McGroarty 1983] use an analytic approach to rating language performance by specifying in considerable detail the elements of language to be observed and how they are to be assessed. The newly developed International English Language Testing System uses different approaches to relate language performance to descriptive scales for each macroskill: the Speaking test uses direct matching of behaviour and global descriptors, the Writing test uses more analytic sub-scales converted statistically to an overall Band Scale, and the Listening and Reading tests statistically convert raw scores on comprehension tests to Band Scale scores [see Ingram 1990, 1990a].

Tests designed to assess some notion of proficiency not only differ on the "directness" with which they approach the task of assessing proficiency but also on the basis of construction of the test. In "Assessing Proficiency", a distinction is made between non-developmental and developmental tests. In non-developmental tests, proficiency is viewed more or less statically as facility in handling certain language content which may be chosen for inclusion in the test on an ad hoc, a linguistic, or a behavioural basis. In contrast, developmental tests see proficiency as developing in an identifiable way from zero to native-like ability levels and they are designed to identify at what point along the continuum from zero to native-like the learner's proficiency falls. Proficiency scales such as the ASLPR, the IELTS Band Scales, the ACTFL Guidelines, the FSI Scale, and the ILR Scale all essentially assert that it is possible to describe how proficiency develops and to slot learners into this continuum. This approach has particular advantages since, as in the ASLPR for example, proficiency is assessed by identifying the sort of language behaviour that the learner manifests, the tasks that can be carried out and how they are carried out, and the point on the scale that most nearly matches the observed behaviour; but such an assessment also indicates the point to which the learner's language has developed and, hence, the direction of future development. Clearly this has important implications for language teaching and enables proficiency scales to make a significant contribution to the teaching process, especially in syllabus design, methodology and the selection of materials and activities [see Ingram 1986].

It was noted earlier that the concept of proficiency lies outside any particular course that learners might follow. The last point just made, however, also suggests that the concept of proficiency is not language specific, i.e., the literature of developmental psycholinguistics clearly suggests that the broad schedule of development is universal even though details of development, the precise ordering of some syntactic forms, for instance, might differ to some extent. Intuitively, we recognize this since we have at least some vague notion of what we mean if we say that Martha is proficient in Thai and Tom has just enough Motu to survive in the village. It is also suggested in practice by the fact that scales such as the FSI, the ILR, ASLPR and ACTFL are successfully used to rate the proficiency of

learners of any language. This observation is important, however, because it makes the development of proficiency scales for the wide range of languages necessary in Australia more practical since it suggests that one does not need a proliferation of scales but rather a common, broad scale which may also include those features distinctive to particular languages. Thus, for instance, the ASLPR is used to rate proficiency in any language, it has been formally assessed and reviewed in the context of ESL, French, Italian and Japanese, the General Description column is common to all versions, and an example column is used to illustrate how those general descriptions relate to particular languages [see Ingram and Wylie 1979/1985 and the versions for French, Italian and Japanese]. For the immediate purposes, this also means that any scales developed for the assessment of Japanese proficiency will probably carry over significantly to other languages and whatever other tests are developed may also serve as useful paradigms for other languages.

In this very brief description of the different sorts of tests that occur and their basis of construction, it is clear, I hope, (and is made clear in "Assessing Proficiency") that different test types are essentially more or less effective, and hence more or less appropriate, for different purposes. Indeed, if one puts aside the possibility of errors in test construction, it is correct to say that there is no intrinsically good or bad test type but rather different test types are more or less effective in achieving different purposes in different contexts; in order to select a test or test type, one must consider the purpose for which one is testing and the context or constraining factors in the test situation and match the test or test type with the purpose and context.

Consequently, in considering what sorts of tests of Japanese may be desirable in Australia, it is essential to consider the purposes for which one requires tests, the needs that testing has to meet, and the contexts in which testing has to take place.

IV NEEDS THAT JAPANESE TESTS AND SCALES MUST MEET IN AUSTRALIA: The needs for Japanese tests in Australia are wide-ranging and, when test construction is to be undertaken, should be considered in some detail. Here we shall only attempt to get an overview of the wide range of needs that exist and the implications they have for test design. In the next three days, these and other needs should be clarified and the test implications elaborated.

Battery: It is already evident that there is a variety of purposes which tests of Japanese have to serve and the different needs will require different test types. It may well be, therefore, that the outcome of this present project should not be just a new test or proficiency scale but proposals for a battery of tests for use with learners of Japanese. How many of the needs can be met and the tests or scales developed will depend on the resources available.

Proficiency: The preceding discussion suggests that, in order to assess proficiency, a direct assessment procedure is desirable but, in fact, this will not always be possible. The initial need is undoubtedly for a proficiency scale, for interview and other elicitation procedures to be worked out, and for alternatives to interviewing to be proposed where it is not possible to use a fully trained native speaking interviewer. In other words, there is need for a proficiency scale or scales in Japanese to be

developed but also for the contexts in which it or they will be used to be considered and a range of application methods to be developed. It may be, for example, that, in addition to standard interview procedures, it would be useful to consider the approach adopted by the International English Language Testing System (IELTS) which uses a wide range of test types and includes a controlled Speaking interview designed to be used by minimally trained teachers; it is possible that the IELTS provides a model that could be adapted for use with Japanese proficiency scales by non-native speaking teachers of Japanese. Two papers describing the IELTS approach have been included in the portfolio of papers for this conference. For Speaking also, it might be desirable to consider whether a "simulated oral proficiency interview" would be useful partly because of the relatively low proficiency levels of Australian teachers of Japanese and because persons outside the education system who need Japanese may have little access to adequately proficient and trained assessors. A short paper on SOPI's by Charles Stansfield has been included in the portfolio.

Whatever other tests are needed and whatever the approach adopted, a proficiency scale or scales should be considered an essential means of explicating and giving coherence to other tests as well as for other, especially course-related, reasons to be discussed subsequently. One of the most significant advantages of the development and widespread acceptance of a proficiency rating scale is that it provides a way to speak in concrete terms about the abstract concept of proficiency [cf. Bachman and Savignon 1986: 380]: it provides common terminology and a means by which teachers, syllabus writers, employers and others can speak more comprehendingly of proficiency rather than by using such vague and quite meaningless terms as "beginners", "average" or "advanced".

Any scale or scales developed will need to take account of the four macroskills and the need to have a different sub-scale for each macroskill. The basic scales would seek to describe general proficiency, they would be developmentally-based, and they would provide descriptors of language behaviour in each macroskill in Japanese as proficiency develops from zero to native-like. Development of suitable scales in Reading and Writing will be especially difficult (but no less desirable) because of the complexities of the Japanese orthography.

In addition, there is need for "special purpose" versions of any Japanese scales developed to facilitate assessment in special purpose contexts, e.g., in Japanese for airline booking clerks, Japanese for scuba-diver trainers, Japanese for tour guides, Japanese for marketing managers, and so on. Desirably these scales would be similar to the general proficiency scale but with adaptations that exemplify the scale in the particular special purpose contexts. The special purpose versions of the ASLPR currently being developed may illustrate what is required [cf. also Biersteker 1986, Doyle 1988].

One "special purpose" proficiency scale of particular importance is Japanese for Academic Purposes for use with students going to study in Japan. At present the only test available is the seriously inadequate Japanese Language Proficiency Test which, it has to be said, is unlikely to make any useful assessment of a foreign language learner's practical proficiency in Japanese or his or her ability to cope at a Japanese institution. In developing a suitable test for use with intending students in Japan, it may be that this conference should consider whether or not an initiative comparable to that which led to the development of the International English Language Testing System is not possible [see Ingram

1990]. The IELTS probably provides a useful model of a test which is designed to provide information on candidates' general proficiency and their "academic purpose" proficiency and is designed to be administered worldwide in situations where there is minimal control over the quality of administrators.

It should also be noted that the need for proficiency scales extends throughout the education system and beyond it. The problem of providing proficiency scales for use where young children are learning a language is particularly difficult and fraught with theoretical and practical problems. The most difficult problem yet to receive adequate research attention is whether or not a single scale can serve for children and adults alike or whether, because scales essentially provide levels within the developmental span from zero to native-like, it is necessary to have one or more separate scales for children. However, as was the case at the end of 1989 when the national curriculum projects in Japanese, Chinese and Indonesian were reviewed, there is an almost desperate need for proficiency scales suitable for use with children throughout the education system in order to assess their proficiency in the language, to explicate the national curricula, to match activities and materials to the learners' levels, and, in many instances, to enable the national curricula to be reinterpreted in some sensible and pedagogically defensible way.

A paper by Elaine Wylie and the present writer on writing proficiency scales is contained in the portfolio of papers. The scales could either be an adaptation of existing scales or be independently developed but should, in any case for reasons of compatibility and communication, be able to be related to the levels of one or more of the established and widely used scales. The development of the scales should be undertaken by a group including native speaking teachers of Japanese as a Second or Foreign Language, native English speaking teachers of Japanese with a high level of proficiency in Japanese, and one or more experts in proficiency and its assessment. The scales should be both informally and formally trialled, presumably in much the same way as the ASLPR was trialled [see Ingram 1984 in the annotated bibliography].

At the time when preparation of this paper was about to commence, the Chairman of one of the language committees of NAATI (the National Accreditation Authority for Translators and Interpreters) suggested that there was no need for work on new proficiency assessment procedures in Indonesian, Japanese or any other language since the NAATI scales and assessment procedures were suitable for the purpose. Such a view misrepresents not only the nature of language proficiency but also the nature of interpreting and translating and suggests a lack of familiarity with the aims and methods of language testing. Indeed, the proposal that NAATI was the appropriate place for the location of some sort of authority for proficiency assessment seems to suggest a view that proficiency assessment is best carried out through translation exercises and, in any case, that translation and interpreting do not involve any skill beyond language proficiency. On the contrary, interpreting and translating are special skills which undoubtedly require proficiency but, in addition, require distinctive interpreting and translating skills, complemented in the profession by appropriate ethical and other work practices. The NAATI scales necessarily take the specific interpreting and translating skills into account and, even if in every other way they were satisfactory as descriptions of proficiency, the inclusion of these interpreting and translating dimensions invalidate them as proficiency scales. The proposal that NAATI should become the administrative authority for language

proficiency tests is unsatisfactory also since, first, if there are persons with expertise in language testing involved with NAATI, they do not publish in the area and, second, proficiency assessment is only one aspect of the array of language tests that are required and it sits more appropriately with the rest of the field of language testing than it does in the specialist field of interpreting and translating.

Industry: There are many uses for tests of Japanese proficiency in industry and commerce. In vocational situations, it is generally not an employee's formal knowledge of the language that is important but whether the applicant has sufficient proficiency to be able to carry out certain tasks (e.g., translate or gloss a letter from a Japanese agent, prepare promotional literature, talk to prospective clients in Japanese, and so on). The precise needs and the full range of needs will have to be identified by the representatives from industry present at this meeting but, in addition, as work proceeds on developing necessary tests, a proper needs analysis may be required. This undoubtedly would reinforce the need for proficiency scales but may also identify the need for other test types.

In industry, proficiency scales serve at least two sets of purposes: on the one hand, they can be used to identify the sorts of skills that employees require in order to carry out particular tasks and, on the other hand, they can be used to measure the proficiency of those employees in order to assess their suitability for employment, to assess their suitability for carrying out specified tasks when they are employed, or to assess eligibility for salary incentives or productivity payments for language skills. For these purposes, either general proficiency or special purpose scales and tests may be required and though common widely used instruments may be suitable in many cases, there will undoubtedly be instances where a specially designed scale, set of criteria or test is required. The source for such assistance should undoubtedly be one of the testing units of the National Languages Institute of Australia located in Melbourne and Griffith Universities or an accredited language assessor as discussed subsequently.

Language Programmes: Proficiency scales are not only useful in the assessment of learners' practical language skills but they have an essential contribution to make to the development, teaching and management of language programmes. This is now a matter for wide-ranging research and discussion worldwide, as some of the papers listed in the bibliographies indicate [e.g. Clark 1983] and the present writer has written at some length on the topic over the last decade [for example, see Ingram 1986, 1985, 1984, 1979/1982]. Without going into detail, it suffices to observe that proficiency scales, because they trace the development of proficiency from zero to native-like, provide an overarching framework within which language programmes can be planned; they enable entry levels and goals to be specified precisely, identify dominant developmental trends that teaching should seek to facilitate, provide a means of assessing the suitability of activities to be proposed in a course, and enable materials to be classified and matched to the learners' stage of development. As already noted and as discussed in the present writer's report of the evaluation of the three national curriculum projects in Indonesian, Chinese and Japanese, those projects and the Australian Language Levels Project have suffered seriously from the lack of proficiency scales suitable for learners in both Primary and Secondary Schools.

Not least, the existence of a suitable proficiency rating system enables the goals to be pursued by courses to be made explicit and the success of courses in developing proficiency to be monitored. At present there is no

information available on what proficiency levels students achieve in their tertiary language programmes in Australia, courses tend to meander along with exercises that may or may not be appropriate and may or may not contribute to the real development of language proficiency, and there are few if any attempts to define graduation requirements in terms of proficiency attained rather than through "time serving", i.e. doing a set number of units. Not only is this unsatisfactory in terms of not knowing what skills learners have on completion of their course, but it allows ineffective teaching and poor course design to go unchallenged.

Quality Control: One of the vital reasons to have satisfactory proficiency scales available and used in language programmes is to exert a quality control mechanism over courses and to increase accountability amongst programme planners and teachers [cf. Bachman and Savignon 1986: 380]. Accountability demands that programmes be so designed with respect to content, methods and time allocation as to enable the specified goals to be attained but if, as just noted, goals do not state unambiguously the proficiency levels to be sought and, if student assessment does not indicate the levels actually being attained, ineffective courses readily go unchecked. In addition, the end-users of the skills developed by courses (i.e., the future employers) can legitimately complain that new graduates lack the skills needed to be of use to them and this in turn puts the whole notion of language skills being vocationally relevant into question.

If quality is to be assured, then appropriate levels of language skills and professional training need to be insisted upon amongst language teachers. For this purpose, appropriate proficiency scales are needed and the minimum skills to be expected need to be specified and written into such controls as teacher registration provisions. In the recommendations for a new State language education policy in Queensland recently prepared by the Centre for Applied Linguistics and Languages at Griffith University, Glyn John and I have recommended that the minimum proficiency for teachers be set at S:3, L:3, W:3, R:3 except that, in character-based languages, it should be not less than S:3, L:3, W:2, R:2.

The fundamental reason for the need for proficiency assessment instruments to ensure quality control is that, without them, it is impossible to obtain and maintain the data on student attainment, course goals, and so on, that are needed if quality is to be assessed: without such scales, essentially, there is no means by which to talk about student performance and no means by which learners themselves can identify their own goals and work towards them systematically. Thus the existence of suitable readily comprehensible statements of proficiency that can be used to specify goals and to talk about proficiency can have a favourable effect on one's ability to monitor the effectiveness of language learning programmes, it can facilitate demands for teacher quality, and it can have a beneficial effect on learners themselves as they become more conscious of what they can do in the language, of what it means to attain proficiency in the language, of the incremental progress they are making, and of the practical usefulness of the sorts of assessment procedures they undergo in order to have their proficiency specified [see also Moeller 1988, Clark 1983, Babcock 1985, Hirsch 1985].

Diversifying and Increasing Language Learning: If the diverse needs in Australia for language skills are to be met and if individual needs and interests are to be catered for, it is essential that the whole structure of language learning be revised. Not least, we have to accept that language learning goes on in other places beside school or university classrooms.

The education system needs also to start to value language skills that may have been acquired other than from formal teaching, e.g. through growing up in a bilingual home, through overseas travel, and so on. If this valuing of other language learning is to be achieved, then at least three things are needed: first, that language courses start to value proficiency attainment; second, that the means be available to assess proficiency whether or not learners have followed a formal course in schools, universities or elsewhere; and, third, that suitable proficiency scales be available and their use built into the regular institutional assessment requirements. In addition, if things other than proficiency are to be valued and built into assessment requirements, then tests for these also need to be devised so that outside learning can be assessed and recognized by the student's institution. Again, the recommendations for a new State language education policy in Queensland include proposals to enable this recognition of outside learning to occur and, thereby, to enable language learning opportunities in all languages to be increased and diversified. Again, however, the key to this is the availability of suitable proficiency rating instruments.

Attainment Tests: The distinction was made earlier between proficiency and course attainment (see also "Assessing Proficiency", Section IV.7). It is an important distinction which is often overlooked by teachers who assume that what they teach in class is automatically going to be available to enhance their students' proficiency or that each item of vocabulary, structure or function learned is going to make a significant difference to the learner's proficiency. It is for this reason that one not infrequently hears the complaint that proficiency instruments are too blunt: this criticism often arises from a failure to distinguish between course attainment or the internalization of what is formally taught and learned in class, from proficiency which exists independently of the courses followed. In that attainment tests are based on the content of the actual courses taught, their construction would seem less relevant to the concerns of the Working Party on the Assessment of Proficiency. However, where students are following courses based on the national curriculum, it would seem advantageous for professionally designed attainment tests to be available both because this would ensure that they were more competently designed and because it would relieve teachers of the necessity of designing their own tests [cf. Clark 1983: 441].

Streaming: Similar needs and answers exist in the area of streaming as for Attainment Testing. Though proficiency may be a factor in streaming, in "Assessing Proficiency" it is argued that course related tests for streaming purposes are probably more appropriate. Where the courses available and into which students are to be streamed are based on a common curriculum such as the national curriculum, there would be advantage in having professionally designed tests available. Thus, for instance, the Stages that are specified in the ALL Project material and in the national curricula in Asian languages (including Japanese) might be accompanied by stage-based tests that can be used for streaming purposes (though, as has been noted in reviews of the ALL Project and of the national curricula, the specification of the Stages is such that it is difficult to know what proficiency levels students entering the Stage will have.

Diagnosis: Diagnostic testing aims to identify the strengths and weaknesses in a learner's language so that weaknesses identified can become the focus of specific intervention by the teacher or learner. To some extent, diagnostic tests are related to the actual content of a course but insofar

as the students are following courses based on the national curriculum and insofar as one accepts that the broad schedule of development in a language is more or less universal, there would seem some advantage in having available a battery of diagnostic tests that are professionally constructed and assist teachers in the course of their regular classwork to diagnose learners' strengths and weaknesses as a basis for remedial intervention. Such tests would need to cover all those aspects of the language and its use that occur in language programmes including the language elements of syntax, vocabulary, cohesion, etc and functional features [cf. Clark 1983].

Self-Assessment: One of the benefits that has flowed from the development and use of proficiency scales is that of motivation resulting from students' being more aware of learning goals and of increases in their proficiency and from their increased ability to monitor their own progress. Self-monitoring is particularly important in the context of diversifying language learning opportunities especially in order to include self-access and autonomous learning. For this purpose, it is desirable to develop tests, scales and procedures that can be used for self-assessment. Proficiency scales for use in self-assessment would be based on the scales for teacher use but would be simplified and be accompanied by additional practical guidance (e.g. questionnaires and examples) to assist learners in introspecting their language behaviour. The scales and other tests would desirably be related eventually to national curricula, to distance education courses, and to self-access materials so as to guide learners through available learning materials.

Self-assessment procedures involving both assisted self-rating scales and other test types could also be incorporated into computer software and interactive video for use in self-access centres, in learning at home, and for assessment purposes in industry. Such procedures could involve all four macroskills, both general and special purpose proficiency, and be interfaced with self-access courses, remedial materials, and distance learning programmes so that the learners can both self-assess and self-direct through the available learning materials [cf. Davies 1984: 48].

Classifying Materials and Activities: One of the advantages of having widely accepted and readily used proficiency scales available is that they facilitate the classification of library materials and other teaching materials. Because proficiency scales describe language behaviour, they allow one to classify materials according to the language behaviour (i.e. the proficiency level) required to cope with them. The ASLPR is already used for this purpose both in language centres and in other libraries and proficiency scales specifically for Japanese would similarly assist. In addition, since proficiency scales allow one to specify entry levels to courses and their goals and to assess the level of language performance required to cope with materials and activities, they provide a way of assessing the suitability of materials and activities to learners and their stage of development. Thus they have an important role in facilitating good language teaching sensitive to the needs and capacities of the learners. Too often when proficiency scales are not available or are not used, teachers adopt materials and activities inappropriate to the level of their students. The most frequent example of this has been the premature introduction of literary studies into language courses at elementary levels. Other examples exist in the national curricula recently reviewed, where, in some languages (Japanese was less of a problem in this regard), reading and writing activities involving characters were proposed that were not at all compatible with the likely proficiency levels of the learners; likewise, the ALL Project materials in general and the curricula based on

them are notably deficient in relating activities to the level of the learners as a result, in particular, of their failure to acknowledge the pivotal role of proficiency in language programmes or to utilize proficiency scales.

Programme Rationalization: At present, little information is available on how rapidly learners of different ages progress through different proficiency levels in Japanese and other languages. The only tables that relate proficiency levels to course lengths are those published by the United States Foreign Service Institute School of Language Studies and a comparable table from Australian sources now publicly available in the recently released *French in Australia* written by Joe Lo Bianco and Alain Monteil. However, these figures relate to languages learned in circumstances that make it difficult confidently to generalize the expectations to courses in the formal education system. Yet, without such information, programme planning is largely based on guesswork, in general administrators and other budget managers grossly over-estimate what can be achieved, and there is urgent need for more accurate and generalizable data to be accumulated. For it to be accumulated, access is required to nationally accepted proficiency scales and the widespread maintenance of records of entry and exit levels, course lengths and types, and student characteristics. The more rational planning of educational provisions and of language learning programmes in particular that would result from the availability of such data would itself more than justify the investment of time and effort required to develop suitable proficiency scales and to maintain the necessary records. Undoubtedly one of the reasons why this sort of data has not been accumulated in the past has been the fear of teachers that it would be used to make them more accountable for the quality of their programmes and the effectiveness of their teaching and, indeed, when, as academic advisor to the Adult Migrant Education Program some ten years ago, I advocated that records be maintained for using the ASLPR, the suggestion was greeted with hostility. For educational planners and language policy-makers, however, the availability of such data would be of very considerable value and, in fact, because most policy-makers and decision-makers outside of the language teaching profession seem to over-estimate the rate at which proficiency changes and hence under-estimate the resources that should be allocated, I believe that language teachers (especially teachers of languages such as Japanese) have more to gain that to lose form ensuring that this sort of data is accumulated. There would, I believe, be much advantage for all concerned (learners, teachers, educational institutions, the Asian Studies Council and others) if one of the outcomes from this Working Group on Assessment of Proficiency Levels for Students of Japanese was not only that suitable proficiency measurement instruments would be developed but that a viable approach to accumulating the sort of data discussed here should be developed.

Other Tests: Reference has been made to the need for proficiency tests in Japanese, attainment tests and diagnostic tests. However, it would be desirable also to consider whether other types of tests might not also be required. One of the most obvious of these is aptitude tests.

The whole field of language aptitude testing has been relatively neglected since the early 1960's when Carroll developed the MLAT which remains the principal test available. Consequently, the aptitude tests available tend not to be satisfactory in the context of modern approaches to language teaching and learning and are excessively constricted in their concept of what aptitude involves. Probably the outcome of a systematic attempt to develop alternative aptitude instruments would not be a single aptitude

test analogous to the MLAT but a profiling battery that would take into account the traditional factors such as aural acuity, grammatical sensitivity, short term and longer term memory, but also factors such as social and psychological distance, introversion-extroversion, age-related factors, motivation, attitudes, previous language learning experiences, and so on. The need for such tests is evident especially in circumstances where employers in industry or administrators in education are allocating resources to enable employees to re-train it is appropriate, before people are selected for such programmes, that their likely ability to succeed in the time available be assessed so that the available places can be directed towards someone else or the available time and resources adjusted to enable them to succeed.

Other tests that warrant consideration include tests of cross-cultural attitudes and of cultural understanding though this meeting may also identify others.

Research: There has been a vast amount of research into language learning and teaching over many years (indeed for several thousand) years and this has proliferated over the last three decades as applied linguistics assumed its role as a distinctive field. However, most of that research has been based on the teaching of English and to a lesser extent on a few European languages while the distinctive problems of the teaching and learning of Asian languages and of Japanese in particular have received comparatively little attention.

The development of valid and reliable proficiency assessment procedures is a pre-requisite for much research into the teaching of Japanese since the purpose of such research is to enhance the efficiency and effectiveness of the teaching and this implies the possibility of being able to make accurate assessments of proficiency changes achieved by alternative methods.

In addition, the problem of developing suitable test instruments (whether direct, indirect or semi-direct) will raise issues that are distinctive for Japanese at least in comparison with European languages. Not least amongst these are the implications of a very different culture and the different linguistic, paralinguistic and other sociocultural ways in which human relationships and interactions are mediated.

Of particular significance for research is the need to better understand the problems inherent in the teaching and learning of character-based writing systems, how reading and writing can most effectively and efficiently be taught, whether there is, in fact, any value in commencing their teaching using Romanized alphabets as initial teaching alphabets, and whether the teaching of reading and writing can be facilitated by the use of such word-processing packages as *Jiejing*, recently developed by Dr Eric Chappell of Griffith University and already being adopted in many countries. In addition, of course, developments in our understanding of the teaching of reading and writing in character-based languages will have important implications for their testing leading to a re-assessment of the whole range of test types for reading and writing and undoubtedly to the development of new types.

Training in Testing: Of no lesser importance than the development of suitable instruments for testing in Japanese is the need for much better training in language testing amongst teachers of Japanese and, for that matter, all other languages. I started the complete version of the paper

"Assessing Proficiency" which is in the portfolio of papers for this conference, by saying that no activity in language teaching was more widely practised and practised so badly as language testing. Regrettably, the situation is little changed some seven years later not least because, in Australia at least, language testing is generally accorded a few short hours in a teacher education course, there have been no specialist courses in language testing available, few if any graduate students have conducted research in language testing in Australia, and, if the literature searches I conducted were adequate, no one has done so in any aspect of the testing of Japanese. This is not just the fault of language teachers or the training programmes, however, since, to be frank, language testers with their traditional obsession with psychometrics rather than psycholinguistics, methodology and syllabus design, have too often seemed irrelevant to the real processes of language teaching and learning. This is, however, no longer the case, especially since recent interest in the concept of proficiency and its assessment emerged in the 1970's. It is essential that, in future, pre-service and in-service teacher education courses for language teachers give due attention to language testing and perhaps one of the outcomes of this meeting could be recommendations to this effect.

Since the early '80's, I have pointed out in numerous papers (including in "Assessing Proficiency") the need for what I have been calling "accredited language assessors". Even if all language teachers were to receive thorough training in language testing, there would, I believe, still be a need for accredited language assessors, not least in Japanese. The fact of the matter is that everyday scores of decisions are made about the proficiency levels of individuals in many different languages and about the language skills that are needed to carry out particular tasks. At present, these decisions are most often made by untrained laypersons with little formal knowledge either of the nature of language or of language assessment. Yet these decisions are often of very great importance ranging from whether or not a person may be accepted into a language programme to legal responsibility to selection for employment, and so on. There is need for individual speakers of a language, for employers, educational institutions, and courts of law amongst many others to have access to persons who are trained both to assess a person's proficiency in a language and to advise on the proficiency required for various purposes. Just as there is established in many countries including Australia an accreditation system for translators and interpreters, I believe there is such a need for language assessors and for similar reasons: the decisions that are made concerning a person's proficiency are often of far-reaching importance and all persons should have the right to valid and reliable assessment and to an appeal process by which injustices if they occur can be rectified. Such language assessors would not necessarily be language teachers but would be graduates with some training in the linguistic sciences, would have taken at least a three month course in all aspects of language testing, and would have demonstrated their skill in identifying language needs and in all forms of language testing (in particular proficiency assessment). A considerable component of such training would, of course, be specific to the language in question, in this case to Japanese.

V <u>TEST TYPES</u>: The types of test instruments that are required to meet the needs of testing in Japanese, especially of proficiency measurement in Japanese, are already indicated in the outline of needs just provided. In brief, they include amongst others:

* proficiency scales (general, specific, global and analytic)

* "free" and controlled interview and other procedures for the application of proficiency rating scales and the scoring and interpretation of test scores

* test batteries (perhaps analogous to IELTS in English) that enable large scale proficiency assessment in circumstances where one can have less confidence in the training and linguistic and professional competence of the test administrators

* alternatives forms of Speaking tests, e.g. Simulated Oral Proficiency Interviews [cf. Stansfield 1989, Stansfield and Kenyon 1988, Clark 1986]

* training procedures for interviewers and assessors using proficiency scales

* self-assessment instruments and procedures

* attainment tests

* diagnostic tests

* possibly other standardized tests aimed at measuring proficiency or some aspect of it

* item banks (validated by trialling and item response theory and related to proficiency levels) that can be drawn on by teachers, for the construction of more widely used tests, and for inclusion in controlled interview procedures

* tests of specific features of the language, e.g. vocabulary, grammar, functions, cohesion, etc.

* computer based tests

VI <u>TESTS OF JAPANESE</u>: In the bibliographies, I have included the names of whatever significant tests and proficiency scales of Japanese that I was able to find. In addition to these, more general scales such as the FSI Scale, ILR Scale and the ASLPR are available. In 1982, the ASLPR was also provided in versions for French, Italian and Japanese mainly as an exercise in demonstrating that it was equally applicable to languages other than English. However, Elaine Wylie and I were not very confident of either the examples provided by our Japanese informant or that the issues distinctive to Japanese had been adequately handled and we have been reluctant to release the scale for use despite reasonably successful trials. It was included in the portfolio solely to exemplify the possibility of providing a language-specific version of a general scale. Since the listing in the bibliographies also provides a short description of the scales, no further assessment will be provided here except to say that there is clearly great need for extensive test development activity in the area of testing Japanese as a foreign language.

VII THE FUTURE: The principal purpose of this conference is to consider the needs for tests of Japanese, especially in the area of Japanese proficiency assessment though it is not possible, in doing so, to exclude consideration of needs for other tests. In the "Advance Organizer" sheet distributed with the portfolio of papers, I raised a number of questions that might direct our thoughts into the area of needs and possible action and in this paper I have sought to canvass what I, as an applied linguist with interests especially in testing, policy-making and curriculum development, see as areas of needs for Japanese testing. The rest of today is intended to consider developments in this area elsewhere and to hear from representatives from industry (the end-users of the products of our courses and of our test results) what they see as the needs for Japanese proficiency testing; over the next two days, those testing experts who have been invited to this conference are charged with specifying needs, identifying what tests are needed to respond to those needs, and outlining in some detail, costing, and prioritizing possible test development projects. Consideration of the projects should include detail of how the tests are to be developed, what documentation is needed, and what validation procedures are required. The comprehensive development process of the IELTS as outlined in one of the papers in the portfolio may be suggestive in this regard and it is worth noting the stages that were undertaken in which needs were identified, the purpose of the test and sub-tests specified, the types of candidates identified, item types considered, test specifications written, item writing teams established, training programmes for test administrators prepared, and comprehensive trialling for purposes of standardization undertaken.

As test development projects are identified over the next two days, it will be necessary to consider how the projects might be undertaken, what role the National Languages Institute of Australia and, in particular, its test centres might have and what input is possible and desirable from other countries, especially those represented here. It will also be necessary to consider how, in the longer term, the tests developed are to be distributed, managed and administered. Some tests, for example, might be sold or otherwise made available to teachers, institutions or accredited language assessors; other tests might be centrally administered, for example by the NLIA; item banks might be stored, perhaps at one of the test centres of the NLIA, and we might consider the need for the Clearinghouse of the NLIA at the University of Queensland to house a databank of tests and items, a possibility that was recently discussed with Roly Sussex, the Director of that Centre, as an activity that the Testing and Curriculum unit at Griffith University might undertake.

Undoubtedly the activity foreshadowed by this meeting will constitute a major contribution to research in the area of the teaching of Japanese as a Foreign Language and, in particular, in the area of Japanese language testing. It will complement the not inconsiderable research and development activity that has gone on in other areas of testing, not least in that of proficiency in English as a Second or Foreign Language, and it provides an opportunity to relate some of that work to the testing of Japanese and to confront some of the distinctive issues affecting Japanese itself, other Asian languages, and, not least, character-based languages. Granted the economic and political importance of Japan to Australia and the increasingly close ties between the two countries, it is appropriate that such research should assume a significant place in Australian applied linguistic activity.

TOURISM INDUSTRY NEEDS AND PRIORITIES

BOB KINNAIRD

R. T. Kinnaird and Associates Pty. Ltd.

Manly, N.S.W.

Mr. Chairman, distinguished international guests and Australian colleagues, thank you for the opportunity to participate in this important gathering.

I know I speak for the Australian tourism industry generally in congratulating the Asian Studies Council for taking the initiative to organise this event; and for recognising the importance of tourism for Japanese language issues.

Japanese language proficiency testing has been a concern of the industry for some time now. Last year, for example, the industry in conjunction with the Australia Japan Foundation, took some practical first steps towards addressing tourism's needs in this area.

Unfortunately, this project became one of the many casualties of the Australian domestic pilots dispute. I hope it can be given a new lease of life through this conference.

I've been asked to speak on the needs of employers in the tourism industry regarding Japanese language proficiency testing.

Last year, the tourism industry was moved to act on Japanese proficiency testing for three main reasons.

First, employers in many sectors of the industry needed a more reliable guide to the actual Japanese language competence of the small but growing numbers of Australians presenting for employment in tourism. As Qantas and many other operators can testify, prospective recruits claim all sorts of proficiency levels, from ail kinds of educational backgrounds.

But unless the organisation had the resources and commitment to arrange some kind of formal or informal testing (such as Qantas did), the employer is completely at the mercy of the applicant.

Obviously, this would not be a major problem if there was some correlation between actual proficiency and years of language study in Australian institutions. But recruitment drives for Japanese speakers have shown, for the tourism industry, there is no such correlation especially for higher education graduates.

The second reason has to do with the Australian tourism industry's commitment to developing and maintaining the high standards of competence and professionalism needed for international competitiveness.

As outlined below, the industry has just gone through a period of extremely rapid growth. To maintain and increase Australia's market share of international tourism, the industry recognised that training standards generally needed attention. Australia's reliability as a tourist destination, shaken as it was by the pilots dispute, depends on reliable standards as much as reliable domestic air services.

For Japanese language proficiency, this mean two things: the establishment of accurate proficiency standards which reflect the actual language requirements of specific critical occupations, and the use of these standards to drive the expansion of Japanese language education and training "production base".

The third consideration behind the tourism industry's move was consumer protection - the consumers here being both people already employed in the industry and those coming through the education/training system.

The last few years have seen enormous growth in the number of private Japanese language training centres offering all sorts of short courses, many of which are taken by tourism's employees. As well as the fly-by-night operators, there are genuine and reputable organisations in the business. But who is to know what proficiency standard you can reasonably expect after say, 30, 50 or 150 hours, unless some reliable authority tells you?

The same concern, I should add, applies with at least equal force to undergraduates entering a 3 year Japanese major.

I'll come back to some of these key issues, but first some essential facts on the tourism industry, and Japanese tourism in particular.

Japanese Tourism to Australia

Australia's international tourism industry has been one of the great success stories of the 1980's and particularly so for the Japan market.

In 1988 - our bicentenary year - Australia achieved over 2 million international visitors from all countries; and for two consecutive years in the late 1980's, Australia recorded annual growth rates of 25 per cent in international visitor arrivals. As a result, international tourism became our largest single earner of export income. To put this achievement in perspective, Australian tourism's growth rate was more than 5 times the OECD average, and the OECD declared Australia as the world's fastest growing tourist destination.

In the case of Japan market, growth rates have been higher. Between 1987-88, the number of Japanese visitors to Australia increased by a staggering 60 per cent. In 1989, total Japanese arrivals in Australia numbered 350,000.

As a result, by 1989 Japanese visitors constituted 17 per cent of the total Australian inbound tourist market.

These are very impressive figures, but the growth potential over the next 10 years is even more impressive.

The Chairman of the Australian Tourism Commission (the Australian Government's own tourism marketing body) announced last month the Commission's targets for international tourism growth to the year 2000. The ATC has sat a target of 6.5 million visitors annually to Australia, by the year 2000. This is nearly 3 times the total number of visitors achieved in the bumper year of 1988.

In the case of the Japan market, the ATC's target for the turn of the century is 1.65 million visitors, or nearly five times the 1989 figure. If we achieve this target of 1.65 million Japanese visitors, then (based on the ATC's estimates):

- Japan will be the most important single market for

Australia's international tourism industry, contributing over 25 per cent of all visitors, just ahead of 'Other Asia' (22%), with Europe on a 15% market share.

- Japanese tourism will have generated around 70,000 additional jobs
- Japanese visitor expenditure in Australia will have reached $6.1 billion per year by the year 2000.

If you are reeling from these figures and the enormous growth potential and benefits for Australia, then consider this: some authorities regard the ATC's growth targets as on the conservative side! The Garnaut Report, for example, concluded that North Asian markets can support much stronger growth than the ATC's targets.

Japanese Tourism and Demand For Japanese Language Skills

There is no doubt that Japanese tourism has been the most important single factor generating labour market demand for Japanese language skills; and also the most important single factor generating student demand for Japanese language learning in Australia.

The Japanese inbound tourism industry has made this contribution both directly and indirectly. Directly, by stimulating demand for Japanese skills in the wide range of different industry sectors that, in combination, deliver the services of a tourism industry to its customers - in transport, accommodation, Customs/Immigration, tourist facilities, etc., - and in the marketing and development of Australia as a tourist destination.

Indirectly, the tourism industry has made a major contribution to demand (for Japanese) in the financial sector, as capital inflows have been drawn into Australia. The list goes on.

For example, Japanese interest in film, media and other services in Australia is ultimately traceable to Australia's impact through tourism, to a very significant extent.

But returning to the specific issue of tourism and Japanese language skills, the plain fact is that we currently have a huge shortage of Australians with Japanese proficiency, from advanced levels to basic proficiency in this industry.

The best indicator of the size of the demand/supply gap currently is data on the number of Japanese nationals working in the industry, compared to the output of graduates in Japanese from higher education institutions.

The Garnaut Report quotes Immigration data indicating that in 1988/89, over 10,000 Japanese nationals were admitted as temporary residents, specifically as "Staff For Travel Companies"; and a further 6,000 or so Japanese nationals came to Australia as Working Holiday Makers, many to work in the tourism industry.

It now appears that the 10,000 figure supplied by Immigration could be an overstatement, but at this stage we don't know by how much - which is itself a sorry state of affairs.

But let us assume for the moment that the correct figure is only 50% of that given to Garnaut - 5,000 people, admitted to Australia because their Japanese language and cultural skills were deemed ""essential"" for the firm's operations; and assume (justifiably) that their skills include advanced Japanese proficiency.

Compare this figure (5,000) with the estimated output in 1989 of graduates with a major in Japanese - probably less than 300. I say "probably" because we have no better data to go on. And if you recall some of the statistics I mentioned earlier, this demand/supply gap is associated with Japanese visitor arrivals of some 350,000 in 1989 - and Australia is aiming for growth to 5 times that number (1.6 million) by end of this decade.

Clearly there is a national need here for careful labour force planning to effectively manage Japanese tourism growth which, on the face of it, must involve accelerated expansion of the language skills production based The tourism industry recognises this need and will shortly be approaching the Australian Government on a joint approach to the issue.

The need for action here has been further highlighted by recent market research. The Australian Tourism Commission's 1989 Survey of Japanese Visitor Satisfaction found that over 60 per cent of Japanese visitors preferred to deal with Japanese speakers (in shops, hotels, restaurants); and two-thirds of these preferred Japanese-speaking Australians.

And research into the major potential growth segments of the Japanese tourist market has identified the "mature age" group as a significant growth market. Service provision in the Japanese language will be essential, to achieve long term growth.

Implications For Japanese Language Proficiency Testing: Priorities

The top priority must be proficiency test development for employment.

Proficiency standards must be clearly established for critical occupations, and these standards must in turn be the benchmark against which proficiency standards on exit from education /training should be measured. Immigration data are the best guide to identifying these critical occupations.

The reason for nominating this as the top priority is not only to meet critical national needs in tourism. This principle of occupation or employment based proficiency is, as I understand it, the same principle that underlines the NAATI system.

This is also the rationale for the location of the Asian Studies Council within the Ministry of Employment, Education and Training. The Council itself was established partly because of past failures to develop appropriate linkages between Asian studies in the education sector. and the employment sector.

From the perspective of the tourism industry, I believe that the next priority must be proficiency test development for students exiting tertiary and secondary education. These are the Australians who will be coming into the labour market or entering tertiary education, and their proficiency

standards must be known - for their own employment and career opportunities.

As far as timing is concerned, it should be clear from all I have said that the tourism industry believes that we now need some tight deadlines.

In fact, my own view is that in the two priority areas for test development that I've mentioned, there should be substantial progress, if not completion, by the end of 1990. I believe we would all benefit from a nationally administered standardised test applied to 1990 graduates in Japanese from higher education institutions These will be the students who mostly commenced study in 1988, the year that first year Japanese enrollments doubled.

I hope it is not necessary to add that in nominating these two priority areas, other areas should not be neglected. Proficiency testing of your children as they progress through schooling is obviously important to the long term goal of turning out bilingual Australians. It is simply a question of emphasis and priority.

Thank you.

EMPLOYER EXPECTATIONS

ROSLYN McLEOD

Tour Hosts Pty. Ltd.

Sydney, NSW

Employer Awareness of Language in Business:

First it must be realised that within the tourism industry and indeed in other industries as well, the employer has to recognise his need for language competency (at various levels) in business. There are still a number of businesses in Australia who do not recognise either the necessity to provide language skills or the benefit their business will gain from engaging such expertise. This evolves from Australia having been a monolingual society for such a long period. There is, however, a growing awareness of the need to provide language inservice within the travel industry but a lack of local resources to do so.

How does an employer judge the competency in language of their prospective employees in this industry? In the main, industry does not expect to have to judge language ability levels; it just assumes that the person applying for the position will have the fluency of the language they offer to carry out the job.

Employer Awareness of Language Competency:

Language skills in the travel industry are merely a tool of communication and not the job itself - the industry does not want to employ interpreters, it wants to employ skilled persons, one of whose attributes is a foreign language. In this respect, language is a passive skill just like personality, good looks or any of the other extras that make a person more successful at their job.

Employer Expectations from Year 12

In considering the expectations of the employer from Year 12 graduates, it must also be noted that the travel industry is made up of many small businesses who are trying to maintain their overheads to a minimum in order to run viable operations. Therefore I expect many of these businesses would draw on students who do not even complete their Year 12.

The industry expects to be obtaining foreign language resources straight from the school system as well as from trade schools, para-professional and professional levels of education. Clearly the system is currently not servicing the needs of the tourism industry because we cannot obtain language resources domestically, as we did not implement the appropriate language education system 15 years ago to fulfil our current needs. If we do not recognise today the dilemma we are currently experiencing and plan appropriately for the future, then we will not achieve the solution we are looking for. Currently we import our language resources. particularly in Japanese.

Different sections of the tourism industry have different needs:

1. Hospitality, restaurants, etc. require students to either enter the workforce directly from Year 12 for in-house training or else to attend a tertiary institution and gain qualifications in various trades.

2. Travel - airlines and agencies require staff with language skills in the ethnic community to service the domestic market, for airline crew to service both the international and domestic markets.

3. Sales and operations - personnel who need to travel internationally to promote Australian products or buy overseas products.

4. Inbound - my area - this segment of the industry relies heavily on tour guides who speak required foreign languages to provide the necessary tour commentary for international visitors as well as liaising between the visitor and the services and facilities they are utilising. The inbound industry also needs its tour consultant staff to have a good understanding of the international visitors' cultural differences and preferences in order to provide good service and acceptable product/itineraries. Other services in the industry such as duty free shops, tourist attractions, restaurants, etc. profit from better client handling through foreign language staff.

Across all sectors of the industry university graduates are needed in areas such as business administration, marketing, engineering, accounting, etc., and it would serve our industry well if more of these graduates were produced from the secondary school system with language to be used as a communication tool to carry out their other skills.

What are the working languages?

Ethnic languages to serve domestic needs.

Whilst we work under the National Policy of Languages which gives priority to Chinese, Japanese, Indonesian along with six others - Arabic. French, German. Greek. Italian and Spanish, for tourism my perception of the demand is the following order: English, Japanese, Spanish, German, French, Italian, Mandarin and Chinese dialects (particularly from Taiwan, Hong Kong and Singapore), Korean, Scandinavian languages and Finnish. Outside of these languages there is very little demand in general inbound tourism.

What does the future hold?

Should we be preparing our young resources graduating from the secondary education system to provide Arabic and all the Asian languages to meet the increased demand at this stage? There is no doubt that the European languages will be more heavily in demand twenty years hence. By this stage Russian and other Eastern block countries will be a growth sector in the tourism industry as well as Mandarin from Mainland China.

Future Growth:

The inbound tourism industry in Australia has gone through considerable growth in the past three years. This has come about because in 1986 and 1987 the Chernobyl disaster and terrorism in Europe turned away tourism from Europe, and we were fortunate enough to scoop up considerable business because we had a successful marketing campaign in place. Paul Hogan, coupled with other attention Australia was gaining in the fields of sport and entertainment, led to unprecedented growth and culminated in our bicentennial year 1988, which included such drawcards as Expo 88 and a host of international conferences that put Australia on the map at No. 4 in the world convention calendar. The downturn in business in 1989 was predictable as most destinations that host a major event usually experience a short term downturn following their big event. However the unscheduled pilots strike has certainly impacted severely on the recovery course in the second half of 1989. Now realistically looking at the future, we must be aware that the world's focus through until the end of 1992 will be on

Europe with the breaking down of the internal trade barriers, not to mention the breaking down of the Berlin Wall.

Meanwhile, the hyperactivity leading up to 1999 has spurred development of new hotel properties and resorts and we will continue to see growth in this sector. This growth may provide an over supply in the short term but will provide the catalyst needed for our industry to go out and promote tourism in this country. Additionally, each capital city in Australia will have a purpose built convention centre which, with the city convention and visitor bureau, will be actively competing in the world meetings market place to increase Australia's share of this lucrative market segment. Coupled with 100% increase in the capacity of the Qantas fleet and similar additional services that will be introduced by other carriers up until 1994, tourism must grow in leaps and bounds so that we will see a much larger industry in the second half of the 90's requiring a far greater output of foreign language resource from the secondary school system to cater for its labour needs.

Prediction:

If the language and cultural learning issue is not addressed in the primary and secondary education sector then tourism in our country is headed for severe difficulty.

Apart from the uncompetitive edge, we must not neglect to cultivate the attitudes in the general community towards tolerance of difference. Increased visitor levels especially concentrating in popular tourist areas needs careful planning to ensure community acceptance as well as the ability to serve the language and cultural needs of the visitor.

The gap between language teaching and business has to be bridged and this has to be a government initiative.

Summary:

Employers expect that education in this country should provide them with staff equipped to do the job. Employers do not want to overlook or reject Australian resources in preference for importing employees. Clearly it is not in the interest of the Australian employer to have to import but currently if business wishes to succeed in a competitive global market place it must provide the service the client expects.

The employer expects the system to produce foreign language ability from Year 12. It also expects tertiary education to provide the basis of training to enable personnel to enter the workforce for further training and experience in their job. It is essential that language be one of the basic skills and indeed should be promoted as part of the prerequisite criteria for entry into any tourism or international trade related courses.

[The author of this paper was prevented by weather conditions at Canberra Airport from giving a presentation at the conference]

ASSESSMENT OF PROFICIENCY LEVELS - JAPANESE LANGUAGE: RELEVANCE TO THE HOTEL INDUSTRY

J. K. LAMBERT

Australian Hotels Association

Sydney, N.S.W.

The hotel industry in Australia is keenly aware of the effect that the Japanese tourist factor has had on their business, and is constantly looking for and achieving better ways of servicing this tourist. For many hotels, the ideal way of providing a good service to their Japanese guest is through the employment of Japanese speaking staff.

Just a brief look at the statistics indicates the importance of the Japanese market to the Australian hotel industry. Over the last decade, the proportion of tourists originating from Japan has increased dramatically in relation to the total number of international tourists coming to this country. in 1989, a year of mixed fortunes for the tourist industry, the proportion still rose to a figure of 17%. Indications for 1990 indicate the proportion is still rising, with Japanese visitors accounting for almost 19% of all international tourists.

The latest International Visitor Survey indicates the relevance of this to the hotel industry, when in 1988 a staggering 2 million nights were spent by Japanese tourists in a hotel with facilities. The market is there, and it is predicted to continue its stunning growth - around 9% in 1990 and 15% in 1991, when the effects of the pilots strike should largely be behind us.

These figures indicate the importance of the Japanese tourist dollar particularly to the international elements within the hotel industry. In fact some hotels rely on this source of income more than others, with certain resorts in Queensland being prime examples of this.

The market responses of hotels in terms of the provision if services to the Japanese tourist are as varied as the establishments themselves. Each hotel or chain seems to have its own approach, and theories on the requirements of the Japanese tourist vary widely. On the one hand, some hotels are providing special liaison officers as well as Japanese speaking staff particularly on the front desk as well as basic Japanese courses for all staff to attend. On the other hand, the theory of some is to limit the number of staff involved in Japanese language on the basis that unless the staff member is fluent in the language and thoroughly versed on the culture, then more "damage" can be done than good.

However, one theme which does appear to be common is that language alone is not enough. In the world of large hotels, the knowledge of Japanese culture and class hierarchy is essential if not to offend the Japanese guest by, for example, booking a senior Japanese executive in a room on a lower floor than his more junior colleague. Perhaps this cultural point is outside the scope of this conference, but it is a worthwhile thought to keep in the back of your minds that to produce a student who is commercially useful to the hotel industry, language proficiency is not the only aspect. However, for the moment I will concentrate on the language standards required in hotels, and elaborate on the services I have mentioned.

The Japanese guest liaison officer seems to be the most specialised form of employee utilising Japanese language skills. From feedback I have received, in most instances these liaison officers are former Japanese nationals rather than Australians using a second language. The reasons for this seem to be firstly a question of availability of suitably qualified Australians and secondly for the reason mentioned earlier that hotels wish

to ensure the cultural aspects are well understood by their liaison officers.

The language proficiency required of these liaison officers is fluency - in speaking, oral translation and listening. As one hotel put it "an Australian who has studied for 12 months is not good enough". For Australians (in the broadest sense in deference to multiculturalism) to be employed in this specialised role, it would seem it is common that some time has been spent in Japan so that the culture as well as the special nuances of the language are known.

Many of the same points are made of those staff members, for example on the reception desk, who are employed specifically for their Japanese language skills. For many hotels catering for the Japanese market who do not have a specialist liaison officer, these employees with added language skills become the focal point for solving Japanese guests' problems.

The third element in the provision of Japanese language service is the basic course taught to hotel staff in general, usually as part of a International visitor awareness course. This course generally includes basic Japanese words, and appropriately Japanese numerals so that the guest's room numbers can be correctly translated. The extent to which hotels go to for this course depends on the philosophy they subscribe to as mentioned earlier - whether a little information can be counterproductive. However, in general there is a belief that the basics are useful.

Now that I have explained WHAT service is provided by the international hotels, it is important to discuss HOW they recruit these people, as this has a direct bearing on your conference this week.

Given the total lack of any proficiency rating system for the skills required, at best you could describe the recruiting as ad hoc. From a tourist industry point of view, I would like to reinforce the comments made in the background notes for this conference that there is a "particular need to assess proficiency in speaking and listening" in the tourist and hospitality industry. Written examinations are virtually worthless from a hotel industry point of view - there is no demand for Japanese reading and writing skills in the industry at the moment, although that is not to say that as the Japanese market continues to grow, there will not be in the future.

At the moment, hotels recruiting staff with Japanese language skills have to rely on either language consultants that are used for their Japanese courses, or more commonly, Japanese staff already employed at the establishment. In the absence of a certificate showing the proficiency of the candidate, the hotel typically invites their existing staff member down for a "chat". I think you would agree this is not a satisfactory situation - it can only lead to an irregularity of proficiency standards.

Consultants involved in at least one Sydney hotel have implemented their own 4-stage oral test, ranging from basic numerics and words to fluency in the language. A system similar to this would be ideal if adopted across the board so that hotels could be confident of the standard being recruited. From the tourist industry point of view, a proficiency system in speaking and listening should be a priority, with consultation with the industry as essential part of this.

If I could now elaborate on a point I made earlier about the employment of Japanese nationals as being a common way for hotels to overcome the lack of confidence in language proficiency as well as ensuring the cultural side is taken care of. It is important to note that this situation is unlikely to continue to the same extent for a number of reasons:

as the growth in Japanese tourism continues, the supply of suitable Japanese nationals with the appropriate work permits or residency is likely to become strained; and,

research indicates that Japanese tourists, while not wanting to visit a culture where everybody speaks their language, are most impressed when Australians, again in their broad sense, can communicate with them fluently

These prospects make it even more important that a proficiency scale is adopted for Japanese language skills taught in Australian educational institutions - and one which suits the particular requirements of the tourist industry. There must be an emphasis on practical language proficiency rather than technical knowledge. Although I am not a language specialist in touch with the latest jargon, it would seem that a special purpose test, or series of tests. could be developed for the hospitality or more broadly tourist industry. However. some care would need to be taken that such tests were not too specific, as it would really be stretching the truth to say that problems arising for guests in hotels were usual, or normal or even common.

Hopefully, today I have given you some insight into the demands of the hotel industry. As the Japanese tourist market continues to grow, so will the demand for quality staff trained in Japanese language skills to an appropriate standard. The hospitality industry overall is moving towards accreditation and commonality of minimum standards in all of its training, and we would like to see that extended to the area of languages, particularly Japanese. The system at the moment is not conducive to either good standards or the employment of Australians using a second language. This situation will only change with the assistance from the academic world to provide the standards and the appropriate training for the industry.

[The author of this paper was prevented by weather conditions at Canberra Airport from giving a presentation at the conference]

PROFICIENCY LEVELS FOR JAPANESE SPEAKERS - QANTAS PERSPECTIVE

SHARON O'DONNELL

Qantas Airways Ltd.

Sydney, N.S.W.

Qantas has developed five specific language proficiency tests to meet its needs. These tests are designed to assess the language ability of Qantas employees and prospective employees according to one of two levels, elementary and advanced.

There is

(1) A General Language Proficiency Test

(2) Language Proficiency Testing - Reservations Sales Agent- Reservations

(3) Language Proficiency Testing - Passenger Sales Agent - Travel Centres

(4) Language Proficiency Testing - Passenger Agent - Airports

(5) Language Proficiency Testing - Cabin Crew.

The general test is designed to assess the candidate's ability to communicate in the language being tested without at this stage possessing job specific vocabulary knowledge. All other tests are designed to assess the employer's ability to communicate in the language being tested with Qantas customers in the performance of their normal job function.

Qantas employees who pass the test may qualify for payment of a language skills allowance - elementary, advanced.

It is also used as a guide to matching employees to tasks and in staff recruitment.

Specifically, Qantas tests the candidates ability in

pronunciation
grammatical structure
comprehension
vocabulary
fluency in speed and naturalness of speech
level of politeness (as appropriate to language and situation).

The candidate is rated on a scale of 0-6, 0 being insufficient ability to pass at elementary level, 1 being elementary pass and 6 being advanced distinction.

The levels are based on the Australian Second Language Proficiency Rating Scale. Qantas needs a test that differentiates at low proficiency levels as well as high. The current scale is adequately meeting this need. Possibly a need to differentiate an intermediate level in the future may be required. The existing scale should meet this.

The general test consists of having the candidates discuss a number of topics which are adapted to the level of their ability.

The specific purpose tests are based on situational dialogues or role plays in which the examiner assumes the role of the passenger and the candidate performs his/her regular job.

At present all tests are conducted through a live interview.

Qantas is interested in a test that is least anxiety producing and structured to elicit language that is as natural as possible. I believe only the general test could be administered through a recorded interview.

Qantas would value a standardised proficiency scale for Japanese speakers and accredited language assessors to test this. The test should assess proficiency in speaking and listening to meet needs.

Appendix

QANTAS LANGUAGE PROFICIENCY LEVELS

The examiner should consider the candidate's ability in each of the following areas when making an assessment:

A) Pronunciation

B) Grammatical structure

C) Comprehension

D) Vocabulary

E) Fluency in speed and naturalness of speech

F) Level of politeness (as appropriate to language and situation).

The examiner is to rate the candidate on a scale of 1-6 as outlined below. (These levels are based on the Australian Second Language Proficiency Rating scale.)

Level 1: Elementary Pass

Able to satisfy basic survival needs and minimum courtesy requirements. In areas of immediate need or on very familiar topics can ask and answer simple questions, initiate and respond to simple statements; and can maintain very simple conversations. Hesitations, fractured sentence construction or many repetitions would be evident.

Level 2: Elementary - Credit

Able to satisfy all survival needs and limited social needs. Developing flexibility in a range of circumstances and can cope with less routine situations. Is able to describe and give information with limited precision within familiar vocabulary range. In work situation can communicate most routine needs and basic details of unpredictable occurrences.

Level 3: Elementary - Distinction

Able to satisfy routine social demands and limited work requirements. Can communicate most work requirements but would need help in handling complications or difficulties. Has sufficient speaking vocabulary to express in simple terms most topics pertinent to everyday life and own environment

Level 4: Advanced - Pass

Able to speak with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social and vocational topics. Can handle with confidence most social situations and those work situations relevant to own needs and experience. Can cope with everyday, difficult linguistic situations. Errors rarely interfere with understanding or disturb a native speaker.

Level 5: Advanced - Credit

Able to use the language fluently and accurately on all levels normally pertinent to personal, social, academic and vocational needs. Can convey exact meaning in social and vocational discussions and can participate in any conversation within the range of own experience with a high degree of fluency and precision of vocabulary. Can respond appropriately even in unfamiliar situations and is always easily understood by a native speaker.

Level 6: Advanced - Distinction

Proficiency equivalent to that of a native speaker of the same socio-cultural group. Has complete fluency, accuracy and range in the language such that speech on all levels is fully accepted by native speakers in all its features. May exhibit minor non-native characteristics but these do not inhibit communication or comprehension.

A BRIEF SUMMARY OF THE INTERIM REPORTS OF SYLLABUS FOR THE JAPANESE LANGUAGE PROFICIENCY TEST (1989-90)

KUNIO OKUDA

Hiroshima University

1. Moji-Goi (Hiragana, Katakana, and Kanji)

1.1 Kanji

Level 4 100 Kanji including the 80 Kanji listed
Survey of Elementary Textbooks (first half)(1989)

Level 3 300 Kanji including the 245 Kanji listed
Survey of Elementary Textbooks (second half)(1989)

Level 2 1000 Kanji (1) Survey of Elementary Textbooks = 598
(2) Educational Kanji List (1989) = 1,006
(3) Newspaper Kanji List (1978) 1 - 1,000
(4) Magazine Kanji List (1962) 1 - 1,000

Level 1 2000 Kanji (1) Everyday Use Kanji List (1981) = 1,945
(2) Newspaper Kanji List (1978) 1,001 - 2,000
(3) Magazine Kanji List (1962) 1,001 - 2,000

1.2 Vocabulary words

Level 4 800 words including the 680 words and 21 greetings listed

Level 3 1,500 words including the 1,317 words and 32 greetings listed

Level 2 5,000 words (1) Survey of Elementary Textbooks = 4,500
(2) Nihongo kyooiku no tame no kihon goi tyoosa (1984) = 6,060
(3) Nihonjin no tisikikaisoo ni okeru hanasi kotoba no zittai (1980) = 5,340
(4) Gaikokuzin ryuugakusei no nihongo nooryoku no hyoozyuun... (1982)

Level 1 10,000 words (1) Words selected for Level 2
(2) Bunrui Goi-Hyoo (1964) 32,600
(3) Kookoo kyookasyo no goi tyoosa
(4) Tyuugakkoo kyookasyo no goi tyoosa

(1) For Levels 1 & 2, classes of words to be excluded from the lists will be mentioned.
(2) For Levels 1 & 2, approximately 80% of 10,000 words and 6,000 words will be listed in the final version.

2. Bunpoo-Dokkai (Grammar and Reading Comprehension)

2.1 Grammar

Level 4 A-1 Sentence Patterns and Grammatical Paradigms
A-2 Particles, Demonstratives, Interrogatives, etc.
1. Interrogatives
2. Demonstratives
3. Particles
4. Derivational Suffixes
5. Discontinuous Adverbials
6. Misc.
B Functional meanings of sentence patterns

Level 3 A-1 Sentence Patterns and Grammatical Paradigms
A-2 Particles, Demonstratives, etc.
1. Demonstratives

2. Contracted forms (e.g. tewa → tya)
3. Particles
4. Derivational Suffixes

B Functional meanings of sentence patterns

Level 2 A Grammatical Function Words (190 items)
B Advanced Sentence Constructions and Patterns
C Honorific Expressions
D Mis . - Pragmatics

Level 1 A Advanced Grammatical Function Words (115 items)

(1) Level 1 items may have subtle differences of nuance and may require more time to be able to use properly.
(2) Those which are more formal and literary expressions will be grouped in Level 1.

2.2 Dokkai (Reading Comprehension)
1. Text types
2. Reading Tasks
3. Reading Skills

3. Tyookai (Listening Comprehension)
1. Text types
2. Listening Tasks
Appendix: List of Skills for Listening Comprehension

4. Analyses of the Test
1. Item Analysis for 1988 Test
2. Difficulty index and discrimination index

[This summary paper was accompanied by handouts which gave example of test items and detailed statistical analyses of test results.]

PROFICIENCY TESTING

TOM WHALLEY

British Columbia Ministry of Education

I would like to thank the Asian Studies Council for their invitation to participate in this international working group on proficiency testing of Japanese. I am not certain that I have much to contribute to this endeavour since the list of proficiency tests of Japanese in Canada is largely non existent. Those that do exist can be described as discrete point grammar tests with substantive amounts of translation included.

To understand this current situation, it is helpful to sketch the background of Japanese language training in Canada. Prior to the start of the British Columbia Government's Pacific Rim Education Initiatives Program in 1987, Japanese was taught almost exclusively at universities and in "heritage" language schools. In universities, the language was usually taken as part of an Asian Studies degree and the application of language skills was usually assumed to be for scholarly purposes. Not surprisingly, tests that were designed for students in this environment were normally associated with reading and writing ability. The writing component was usually very insignificant. Predominantly what universities had then were reading tests to determine proficiency. These exams, even as reading exams, were also quite limited in sophistication compared to reading exams one might find currently in E.S.L. The heavy inroads of the audio lingual method with its emphasis on rote learning of hierarchically ordered sentence structures tended to mould reading exams so they focused on reading passages designed to present those structures; little thought was given to anything but linguistic features in the choice of text.

In the heritage language schools, the emphasis was also likely to be on reading and writing since many of the children were near native speakers of Japanese. The shape of the proficiency testing experienced by testers in Japan was really what guided the kinds of tests that were put in place in the heritage language schools. They can be described as tests of discrete items - vocabulary, sentence structure and grammar.

To get an idea where we currently are and where we might want to go, I would like to examine a recently developed Japanese proficiency test developed in B.C. Let's start with looking at the instructions for each part of the test. Each section has instructions as follows:

1. Look at the pictures and write one word in HIRAGANA.
2. Look at the pictures and write one word in KATAKANA.
3. Choose the correct reading for each KANJI.
4. Write KANJI for the underlined words.
5. Select the correct answer for each.
6. Select the correct word for each blank.
7. Select the correct word for each sentence.
8. Read the following and answer the questions.

Numbers 1 and 2 cue children with pictures of common items chosen apparently randomly; certainly students would see no logic to the choice. Number 3 gives a seemingly random list of single Kanji and gives a choice of four "kun" readings to select among. The distractors have some association with the character. Number 4 uses Hiragana to supply the cue for the Kanji. Number 5 presents questions and answers in short dialogue and asks the student to choose the correct response from among four choices. Number 6 asks the students to insert the correct particle from a choice of four. Number 7 asks students to select the correct word from among four presented. These words are groups of nouns in some instances adjectives, adverbs, conjunctions or pronouns in others. Number 8, the

final question, requires the children to read a short passage containing simple sentences related in a narrative.

We can make a number of observations about this test. First, it is a test of USE not USAGE; it requires a theoretical or analytical knowledge of the language. The student is not asked to do anything under specific conditions with tasks requiring the language to do something. Second, this is a NORM REFERENCED TEST not A CRITERION REFERENCED TEST; it will be used to describe how well one student does against another student or a numerical standard. This test will not tell you how well a student can do a specific task such as determine the age and sex and hobbies of a Japanese peer from a "penpal wanted" advertisement. Finally, we can say that this test uses text that is UNAUTHENTIC rather than AUTHENTIC; the student is unlikely to ever see Japanese as it is in the tests in this way except on a test. The student is not exposed to any document he may see in the "real" world nor is he asked to do anything that he may be required to do with the language. Finally, we can observe that this test primarily tests READING and WRITING SKILLS, not AURAL/ORAL SKILLS; there is, in fact, such a component to this test, but it is not very developed and is characterized by use, norm referencing and inauthenticity.

What is wrong with this test I have described? In short, it does not do the job. It is not a proficiency test. I don't think that it tells us much about a student's communicative ability with the language; rather, it measures the student's analytical or theoretical knowledge of the language. It would be very hard to tell someone what a student may be capable of doing with the language. Can he read a postcard? Can he follow directions? Can he recognize warning signs, exits, entrances, etc? Moreover, it is very difficult to imagine how the description of one level of this test can vary from the next. Each subsequent level requires the student to simply do more; write more words in Kana, recognize more Kanji. The test is not a very sophisticated instrument. Perhaps it is intended to measure broad and generic language skills but it yields a description that may be useful for placement but seems of little use for measuring proficiency.

It seems too clear that if we are aiming to develop proficiency tests that measure what a student can do with the language, we need to design texts that are tests of <u>usage</u>; tests that specify how the student is to use the language. We need to design tests that are <u>criterion referenced</u>; that spell out how successfully the student met a task. And finally, we need to design tests that are authentic; the tests should use authentic formats and authentic activities.

In addition to these broad orientations, we need to address a more specialized agenda to design tests that measure what students can do, their proficiency. In his 1980 publication Testing Communicative Competence; An Interim Report, Brendon Carroll laid out a specification procedure which I think is very useful. I have inserted it for your consideration.

	Information Banks	Procedural Guide
1	Participant identification	Broadly describe typical participant.
2	Purpose for language use	Describe main uses of English and classify under ESP headings: academic, occupational or social survival.
3	Events/activities	Choose the major events to be met with, and select several activities for each.
4	Instrumentality	Select media - listening, speaking, reading or writing, or multiple-mode communications. Channels: face-to-face, tape, print, film, etc.
5	Socio-cultural	Specify social relationships, dialects and socio-cultural factors.
6	Performance levels	Using nine-point scale, give target levels of performance for each medium and multiple-mode.
7	Topic areas	Identify semantic areas for each specified event.
8	Language skills	Choose the skills necessary for carrying out the different activities at given target levels.
9	Language function/tones units	Indicate functions needed, and appropriate attitudinal tones, for those activities involving sizeable person-to-person interactions.
10	Test format	Choose types of item for each activity - closed-ended, open-ended or restricted response (note RACE and authenticity).

[The author concluded his presentation at this point by saying that he felt that he had a better idea of what was not successful than exactly what form and shape tests might take. He hoped, however, that during the days following he would be able to share some of his thoughts about principles that are desirable in proficiency tests. To indicate directions that proficiency testing in the Japanese language in British Columbia is taking, he had with him some examples of tests that he and colleagues had been developing - attempts to assess how well children have reached curriculum goals, for example whether Grade 4 children have achieved one of the outcomes of a unit on food, to express likes and dislikes about foods].

FOREIGN LANGUAGE PROFICIENCY ASSESSMENT IN THE US

DAVID V. HIPLE

American Council on the Teaching of Foreign Languages, Inc.

Yonkers, NY

and

HIROKO KATAOKA

University of Oregon

(summary of presentation)

I. President's Commission on Foreign Language and International Studies, 1979
   A. Major recommendations:
      1. Assess proficiency of students/teachers in foreign language programs
      2. Determine language skills required to function in a variety of contexts
      3. Create multi-level statements of global proficiency, with special attention to speaking proficiency

II. ACTFL Proficiency Guidelines
   A. Characteristics of Proficiency Guidelines
      I. Descriptions say nothing about time or effort involved in reaching a level nor length of time spent at any given level
      2. Descriptions are geared to performance of real life language functions in ever-expanding personal, social, and professional contexts
      3. Descriptions are sequential and each level subsumes all previous levels
      4. Descriptions reflect highest level of sustained performance
      5. Descriptions reflect degrees of control - awareness, partial control, full control

III. ACTFL Oral Proficiency Interview
   A. Characteristics of the OPI
      1. Adaptive
      2. Interactive
      3. Humanistic
      4. Referenced against interlocutor
      5. Global rating includes five assessment criteria
         a. function/task, context, content, text type, accuracy (vocabulary, pronunciation, grammar, fluency, socio-linguistic competence)

IV Interagency Language Roundtable
   A. Compatibility with pre-existing government system
   B. Human resources, the living tradition
   C. Guidance and endorsement
      1. Training program
      2. Standards

V. Language-specific Implementation
   A. Western/Indo-European bias
   B. Asian language comparability
   C. Interlanguage reliability

VI. Language-specific Initiatives
   A. Japanese proficiency guidelines
   B. Japanese tester training
   C. Proficiency-based instruction in Japanese

LANGUAGE PLANNING AND THE TESTING OF JAPANESE IN AUSTRALIA

J.V. NEUSTUPNY

Monash University, Victoria

LANGUAGE PLANNING AND LANGUAGE TESTING

Language testing is closely associated with language teaching. Language teaching, in turn, is a fact of the broader area usually referred to as language policy or language planning. In this paper I wish to depart from the contemporary theory of language planning and proceed to the discussion of how some recent theoretical approaches to language planning (or "language management") can contribute to the issue of language proficiency assessment, central to this conference.

Theories concerned with the treatment of language problems started developing before World War II within the Prague School of Linguistics (Vachek 1966) but the present discussions in the discipline draw mainly on approaches which emerged with the growth of sociolinguistics in the 1960s. Within sociolinguistics, the study of language problems has consolidated as an independent stream, with Einar Haugen, Joshua Fishman, Joan Rubin and Bjorn Jernudd as the principal contributors (cf. Cobarrubias and Fishman, eds. 1983).

In the 1980s some new trends surfaced in the theory, mainly in the work of Joan Rubin, Bjorn Jernudd and myself. While the discipline so far was principally concerned with macro-planning, the new approaches emphasized the importance of tracing language problems in the acts of particular discourse. The theory on which these approaches were based has usually been referred to as the "correction theory" (cf. Neustupny 1977,1978, Ch.12, Rubin 1978- 9, Jernudd & Thuan 1980, Jernudd 1982, 1983, Neustupny 1985a). Rather than considering a possible spelling reform in abstracto, we claimed that language planners should commence by finding out what language problems occurred in the actual acts of reading and writing in the case of particular persons at particular times.

Within the last decade, those working in the theory of language problems have felt uneasy about the term "language planning" (Jernudd & Neustupny 1987). It has been suggested that it should be replaced by the term "language management" which then covers "simple correction" of language in discourse, "language treatment" and a number of other management systems among which "language teaching" occupies the foremost position. In this terminological system the use of the phrase "language planning" is restricted to reference to one particular approach to language treatment and some other management systems, namely the approach which has emanated from the work of Einar Haugen and Joshua Fishman and was characteristic for the 1960s, 70s and 80s (see particularly Rubin & Jernudd 1971).

The general structure of the processes of language management had already been identified, already in the 1960s (Haugen 1966) as consisting of
A. Problem identification
B. Policy selection
C. Policy implementation and
D. Feedback.
Language teaching commences with problem identification (typically the realization of the lack of competence in a language), proceeds to course design, and implements the design in the teaching process. The feedback stage (D) is identical with the stage of problem identification (A) after the initial cycle (A-C) has been completed. We can thus speak of a continuing process of management consisting of a number of cycles: A1, B1, C1, A2, B2, C2, A3, B3, C3 ... However, at the stages A2, A3, etc. a number of sub-disciplines have emerged and assumed relative independence:

one of them is the study of language acquisition, another one is language testing.

1. TESTING AUTHENTIC INTERACTION

In the theory of language management the role of individual micro-processes is emphasized. This is also true of the process of testing. The testing process is based on:
(1) Selecting an appropriate speech sample
(2) Noting of deviances from the target norm (Neustupny 1985a)
(3) Evaluating such deviations and
(4) Scoring.

With regard to the first of these stages, the current testing theories work with either "indirect" or "direct" data created for the purpose in testing situations. The language management theory assumes that all language management should commence with authentic discourse data. The consequence for language testing is that language testing, too, should be based, in the first place, not on "realistic" but on "real" discourse. How does this requirement translate into the practice of language testing?

Some of the speech used for testing is elicited in the minimum of context (indirect testing), while some has a relatively full context (direct tests) - in all cases which came to my attention, these contexts occur in testing situations. However, normal interaction takes place in situations of authentic use. We are interested in how people perform in such authentic interaction, not in testing situations. Therefore, we should be interested in testing that performance directly.

Not only is authentic interaction our aim. We are also increasingly teaching Japanese through authentic interaction. I have previously called this the "Interactive Competence Approach" (Neustupny 1989).

Monash is currently developing two introductory-to-advanced courses (Comprehensive Japanese and Business Japanese) which are based on authentic interaction ("performances") rather than simulation or role play. Each segment of the course is centred around the students' participation in an authentic interaction with Japanese speakers. Another new development is bilingual education in Japanese for primary school children, a program developed presently in Victoria by the Japanese Studies Centre in cooperation with the Victorian Ministry of Education. In a bilingual course students communicate in order to learn physical education, music, maths, etc.; the purpose of learning the language is hidden. This is authentic interaction, not just language teaching.

In an immersion course (Neustupny & Ozaki 1986), such as is a part of the new Monash undergraduate program, in which the students study about the Japanese system of education through the medium of Japanese - not just learning Japanese using the topic of education - a number of cases of authentic interaction ("performances") take place. There are lectures, panel discussions, individual work with statistics, etc. The course is closed with each student interviewing individually for a written assignment a Japanese family about the problems of educating their children in Melbourne. What I wish to claim is that the assessment of the students' proficiency should be based on their performance in these authentic activities, rather than on their performance in testing situations alone.

Sharon Lapkin (1985) describes testing instruments designed for students in French immersion courses in Canada. Her description seems to indicate that such testing is conducted in special testing situations, rather than utilizing the students authentic performance during the course of bilingual instruction.

What is a more valid test of competence in guiding Japanese tourists - a testing interview or the assessment of authentic performance in a real guiding situation?

Authentic interaction testing represents an important addition to the existing typology of tests. It differs radically from the indirect and, to a considerable extent, from the interview-based direct tests. Understandably, there will be doubts. Defendants of the interview paradigm will continue claiming that authenticity cannot be achieved. They will find arguments against the validity, reliability and practicality of authentic tests. I do not wish to claim that these issues are trifle. What I do wish to claim is that we must not stop at the level of the interview testing. Language testing must proceed to the next paradigmatic stage. And this is also true of Japanese language testing.

I assume that the new paradigm in language testing will include authentic testing as far as is practicable. The word include must be emphasized. We cannot expect that all aspects of interactive competence can be tested through the same instrument. Direct testing, particularly of the interview type, will necessarily remain; so will semi-direct and to a lesser degree indirect tests.

Let me add a few further comments on authentic tests. Like any work which concerns interaction problems, they require an assessment of the processes of noting and evaluation. In conversation, not all deviations from norms of the target system are noted, and not all of them are evaluated as inadequate. For instance, there is evidence that only a minority of deviations from the Japanese accentuation are noted by native interlocutors and that only some types are negatively evaluated. This point is missed in the existing testing paradigms. In most indirect tests any deviation is noted by an experienced examiner and negatively evaluated, irrespective of the likelihood of such evaluation by native participants in the discourse. In interview tests, deviations are noted and evaluated by a participant, but the participant is the tester who has been trained in noting and evaluating. He/she is not any more a normal participant.

We need to establish scales of different degrees of inadequacy for different deviations from Japanese norms. It is obvious that in most situations it will be more inadequate to address a senior interlocutor with the word anata "you" than to make a mistake in a particle. This is a basic issue for testing. At present our understanding of this area is close to zero.

In addition, foreign participants in contact situations also note and evaluate features of their own speech. This matter is important for assessing the speaker's learning competence, and I shall return to the point later in this paper.

Authentic tests provide thus an opportunity to assess, for the first time in the history of testing, the degree of seriousness of deviations from the target norm. There are several methods (such as the follow-up interview) but they cannot be discussed here.

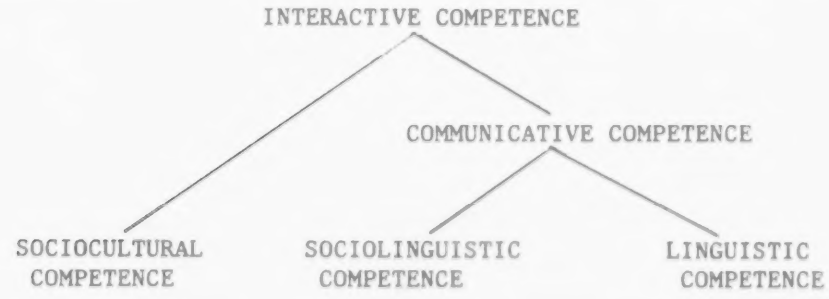
2. TESTING SOCIOLINGUISTIC AND SOCIOCULTURAL COMPETENCE

The Haugen model of language planning worked principally with linguistic competence alone. The management model strongly claims that language problems must be considered along with sociolinguistic communication problems and sociocultural problems (Neustupny 1977). An attempt to solve the problem of illiteracy by teaching reading and writing must fail, unless we can introduce measures that make people use reading and writing and unless we create sociocultural conditions which allow the use of the written language to happen.

In language teaching the importance of "culture" has been paid lip service. There is a number of books entitled "Teaching Culture" or similar, but even today we still lack a simple check-list of items from which a course developer could draw.

Existing testing theories ignore sociolinguistic and sociocultural competence almost completely. Introductions to language testing such as Testing for Language Teachers by Arthur Hughes (Cambridge 1989) contain no reference to anything other than language testing. The first weak indication that a change might be coming are the current American discussions recorded in Thompson, Thompson & Hiple's paper (1988). My literature survey, though admittedly limited in extent, has revealed a paper by Rebecca Valette (1977) entitled The Culture Test. It identifies four areas of "culture" that deserve testing: greater awareness and broader knowledge of culture, etiquette, differences between the target and the student's culture, and the values of the target culture. Surely, there must be more that requires attention.

One of the first problems is that some basic distinctions are ignored. Firstly, within the are of "culture" we must distinguish between a vast area of non-linguistic communicative rules (I shall call them sociolinguistic rules) and non-communicative sociocultural rules. (Sometimes I use the word "socioeconomic" to emphasize that the range extends to economy as well. The categories to be considered are:



More than enough has been done for the mapping of linguistic competence. The situation is less clear with regard to sociolinguistic and sociocultural competence. Let us examine each of them in turn.

2.1. Sociolinguistic Competence

An excellent model for sociolinguistic competence has been provided by Dell Hymes (1962, 1972). The problem here is that it is poorly known. In my own adaptation of Hymes' framework I speak of:

1. Switch-on rules (under what conditions interaction occurs)
2. Variety rules (which "kind" of language is used)
3. Setting rules (when and where interaction occurs)
4. Participant rules (who are the people who participate in interaction)
5. Content rules (what is the content of interaction)
6. Message rules (in what form interaction takes place)
7. Channel rules (through what medium; this category includes what is normally called non-verbal communication) and
8. Management rules (which label, evaluate and correct interaction acts).

These categories of rules operate in situations, which are conceived as units of the process of interaction. The Hymesian model thus implies two basic procedures: the listing of situations and, within each situation, the listing of the rules of interaction (1-8 above) applied.

One important point when we think of communicative situations is to recognize that apart from native situations (situations in which all factors, such as participants, belong to the same culture) we also have to deal with contact situations (in which some factors, e.g. participants, are foreign; cf.Neustupny 1985a). Language teaching and therefore also language testing by definition, deal with contact situations alone. We are not interested in how the Japanese behave when they are among themselves, but how they and we behave, and should behave, in situations of contact. The issue of norms of contact situations is a complicated one (Neustupny 1985b, Marriott forthcominq 1) but it must be approached before we attempt any realistic form of testing. For example, a wide range of English loan words is normally acceptable in Japanese but not in their English phonological forms. They must be pronounced in Japanese. The use of the title -san is acceptable in contact situations conducted in English, but not in reference to persons who would receive a higher title in Japanese. A foreign participant may be excused for not assuming the formal sitting posture on tatami but may not be excused if he sits in a very informal way on a chair.

A considerable amount of research, using this model for the study of Australian-Japanese interaction, has been conducted at Monash University (Neustupny 1988). The results of this research are summarized in my book Communicating with the Japanese (The Japan Times 1987). In particular chapter III of the book can be used directly for the construction of proficiency tests designed to cover sociolinguistic competence.

Let me give some examples of what is involved: speech may not be required in some situations (such as meal situations), the selection of the proper variety (e.g., Standard or Common Japanese) is important for achieving close personal contact, the selection and manipulation of honorific styles determines the relationship between participants, the structure of appointments (time schedules for communication) differs, the communication focus (address) for academics is their home, not their office, the use of honorifics in middle class situations is vital even at the beginners' level, a vast number of politeness rules affect language as well as speech (the whole problem of etiquette belongs here), the use of pronouns is

strongly constrained, titles are obligatory in some contexts, the form of correct requests is considerably complicated, rules for developing friendship networks must be mastered, the structure of parties requires proper attention, without proper "back-channelling" speakers may be assessed as uninterested in interaction, shop attendants do not normally establish semi-private communication networks with customers, there are good and bad topics, it is normally inappropriate to admit to one's good knowledge, it is normally inappropriate to communicate one's strong likes or dislikes, the communication of approval and disapproval is indirect (though it does take place), letter structure differs, a large number of opening and closing routines for situations (greetings, etc.) must be mastered, the use of greetings (hello, good evening, good bye) is mostly radically different from English, the written channel of communication has a basic importance, gestures are socially constrained, laughing and smiling needs considerable attention, so does nose blowing, bowing, sitting and standing posture, formality of dress, etc., gift giving is an inescapable procedure, requests for correction to senior interlocutors may be evaluated negatively, etc.

The consequences for testing are many. Yet, no proficiency instrument which I have encountered takes the matter as seriously as it should be taken. Some of the sociolinguistic rules may affect the transmission of the message (e.g. the appointment rules). Many others ""only" support or destroy the presentation of self. For example, if participants fail to receive or to communicate their intention to establish an informal relation, is this not a major failure in their proficiency in Japanese? Obviously, the relevance of sociolinguistic rules of communication for Japanese interaction proficiency is fundamental.

Some of the sociolinguistic rules are currently taught and their knowledge is tested. For example, the differentiation in the use of greetings depending on participants and their relationship, honorific switching, replies to compliments, ("your Japanese is excellent") etc. Further development is urgently needed.

2.2. Sociocultural Competence

The mapping of sociocultural competence is more complicated than that of the other two. Anthropologists and sociologists have failed so far to produce for us what we would call an "explicit grammar of culture and society". The best way to proceed is to use the concept of domain of interaction as established by Fishman (1972). A domain consists of situations that share common features. The concept lacks theoretical elaboration but it works well in practice. The principal domains are:

1. Daily life domain
2. Family domain
3. Friendship domain
4. Education domain
5. Work domain
7. Public life domain
8. Services domain
9. Entertainment and culture domain.

Obviously, this is an "etic" (universal) list which must be converted into an "emic" system (Befu 1989) that represents the particular situation within the Japanese culture. However, even in its etic form the system is helpful.

Each domain can be broken down into situations. It is usually relatively easy to map sociocultural components of behaviour within a situation and proceed towards more general rules valid within a domain. It is necessary to keep in mind that apart from overt behaviour there is also knowledge and attitudes, which need attention. When interacting with Japanese tourists we must not only know what their social characteristics are but also what covert attitudes they hold towards Australia.

It is difficult to imagine how an Australian could interact successfully with a Japanese in Australia without possessing the basic information on the Japanese community in Australia: how they conduct their daily life, what their family life is, how they establish and maintain friendship, how their children are educated, where and how they work, what their attitudes to entertainment are, etc. How can an Australian businessman successfully communicate with his Japanese counterpart if he does not know what the structure of Japanese institutions in Japan is or how the structure of Japanese institutions (company, Japanese Society, etc.) changes when they are transferred to Australia. Can we communicate with the Japanese if we are not sure what they read, whether they have a religion, etc.

A number of questions arises for testing. While the frequency of linguistic rules is high and we can easily cover them within an interview of 25 minutes, the occurrence of sociolinguistic and sociocultural rules of behaviour requires much more time. Interviews obviously cannot test in a direct way familiarity with personnel or settings of any other situation than the testing situation itself. Yet, the matter of testing sociolinguistic and sociocultural competence is very serious.

In a different context (Neustupny 1989) I have postulated three different types of Asia- and Japan-literacy. The objective of Japan-literacy 1 is the ability to use the understanding of Japanese society and culture as a means of interpreting Australian-Japanese economic, political, social and cultural interaction in a very general way and, perhaps, occasionally interacting with the Japanese (through the medium of English). A considerable deal of sociocultural competence for Australian-Japanese contact situations is required, but only some sociolinguistic competence, and virtually no linguistic competence. The target should be the majority of Australians. In order to encourage the development of this type of literacy and to assess its levels, the development of testing instruments is overdue.

Japan-literacy 2 is required by those who, apart from interaction similar to that covered by Japan-literacy 1, also undertake extensive interpersonal contact with the Japanese through English. This category will include most government and business personnel. Apart from a high level of socioeconomic competence these participants need a considerable deal of sociolinguistic competence necessary to communicate with the Japanese in English. We must get away from the popular belief that international communication through English follows the norms of native English communication. The norms of sociolinguistic behaviour in contact situations which use English can differ substantially (cf. Marriott forthcoming 1). An awareness, knowledge and active competence in behaving in such situations is needed. For Japan-literacy 2, apart from the study of socioeconomic competence, a good grounding in sociolinguistic competence is needed. Only a smattering of linguistic competence will do. It is unrealistic to believe that in the future almost everyone in government and business will speak Japanese.

Again, testing instruments are inevitable.

Finally, Japan-literacy 3 is for the relatively limited number of those who will be working in permanent interpersonal contact with the Japanese. This should include a considerable number of those in government and business. Apart from socioeconomic and sociolinguistic competence they need varying degrees of linguistic competence in Japanese. When I say a limited number of people in this category is needed, I should add that in my view this number must be a multiple of the numbers we have at present. General proficiency and special proficiency tests must cover, in this instance, all three types of competence.

No need exists or will exist for tests which cover only linguistic or only linguistic and sociolinguistic competence. The need for testing instruments dealing with Japan-literacy 1 and 2 is urgent, the need for testing Japan-literacy 3 less so.

3. LANGUAGE TESTING AND LANGUAGE TEACHING: THE BACKWASH

The model of language management which defines language testing as one of the processes of the stage A2 (problem identification of the second cycle) of language teaching, implies that information obtained through testing is available for policy selection. In the case of language teaching this is identical with course design. This fact is of course well known by those working in the area of language testing and has been known as the harmful or beneficial "backwash" on teaching.

The matter is serious. A prestigious test which covers only a part of linguistic competence, for instance only the spoken or only the written language, will necessarily result in redesigning courses to maximize the probability of students' passing the test.

A test covering linguistic competence alone, would strongly endanger those programs that include sociolinguistic competence as well.

Attempts to teach all components of interaction with Japan, rather than simply language and communication, would necessarily suffer should the sociocultural component remain untested. Students who invested their time in the study of Japan, as well as the language, would be disadvantaged in comparison with those who know little about Japan and the Japanese but have used the extra time available to remember additional linguistic patterns. Students who graduated from what I see as the more progressive programs would be losing in competition for scholarships, jobs and in all areas where testing results would be required. The inclusion of an unrepresentative selection of some limited "cultural" facts will not help.

At this moment the only test administered in Australia that has a potentially harmful backwash is the test used for the Mombusho scholarships. It emphasizes written language competence and favours therefore students who, instead of investing in a balanced competence in the spoken and written language, used all their time for the study of grammar and reading.

Another possibility of a harmful backwash is connected with the form of the test. Indirect testing favours programs which place emphasis on the study of rules without a full context. Students who can successfully interact in authentic situations are disadvantaged if this essential ability remains untested.

On the other hand, the backwash may be beneficial. A set of tests which go hand in hand with the recent advances in curriculum development will support such development and push for progress in Japanese language teaching. However, we must be very careful. If the power of testing is very strong or if we apply a test too widely, the result may be harmful to further course development in the area. Backwash is a weapon that must be applied in a highly sensitive way.

Another matter that needs careful consideration is the various functions of Japanese language teaching in Australia. Proficiency tests only test proficiency. In particular secondary school Japanese programs fulfil a number of functions that are not connected with proficiency but are still useful. For example, the teaching of Japanese works as a tool for increasing the understanding of other cultures and as a tool for skill-formation. These are legitimate aims, and should not be negatively affected by the establishment of testing programs.

SOME SPECIAL FEATURES OF JAPANESE TESTING

Language testing theories have concentrated either on the testing of English as a second language or testing of languages which both linguistically and culturally are very close to English. It would be incorrect to say that Japanese has been left out. There is the proficiency test developed in Japan, there are Ingram and Wylie's proficiency ratings, the ACTL oral proficiency interviews. But, on the whole, my impression is that the future of Japanese language testing is still before us.

Therefore, let me note a few points which I feel are important for further work in this area. Some of them are particularly valid for the Australian situation, others are of a more general nature.

1. The standards of Japanese language teaching and testing in Australia

Australia is not only one of the two countries (together with Korea) which have the highest percentage of learners of Japanese per head of population (Neustupny 1975), but also a country with probably the widest selection of courses and the most modern methodology applied in these courses. We are teaching Japanese at the primary level (both in conventional and bilingual courses), we have the most highly developed system of secondary language teaching in the world, with some excellent courses in existence and others in making, and tertiary undergraduate language teaching of considerable extension. I know of no other country that has developed postgraduate language or communication courses such as the Queensland interpreting course or the Monash Business Communication program. Most of these programs use an advanced form of what could be classified as a communicative approach. Some programs, being interactive, go beyond the narrow range of the communicative teaching.

One area in which adjustment is needed is the methodology applied in tertiary courses. Some of the approaches used are still close to the grammar-translation mode, others approximate the audio-lingual paradigm. Especially after the substantial growth in enrolments of the last two years, much corrective work is due in this area.

The second point to be made is that advances in course development have not necessarily been accompanied by parallel advances in testing. This, again, is particularly true of the tertiary sector.

2. Development of JSP

Japanese for special purposes has strongly developed in particular in the last three years. Perhaps the first was the Griffith Science Japanese course followed by the translation/interpreting program at Queensland and the business communication program at Monash. Various tourism courses have been developed, and forthcoming developments include legal Japanese and engineering Japanese.

At least some of these courses would profit from the development of parallel testing instruments. It is important that the work which has already been undertaken in course development and in some cases in research (especially business Japanese and tourism Japanese at Monash, cf. Marriott 1988, 1989 and forthcoming 2) is not ignored and existing expertise fully utilised.

3. The level of achievement and testing

The system of Japanese linguistic competence is very distant from English and other European languages. Particularly difficult are the order of elements, the lexicon, honorifics, discourse structure and the script. Because of this, progress is slower than in European languages.

When testing occurs after the first say 150 or 300 hours of instruction, the competence of the learners still remains limited and different areas may have been covered depending on the course undertaken. For instance, the phrase "the first most common 100 characters (kanji)" makes no sense. Which 100 characters are taught first depends on the course undertaken. Another factor is the orientation of courses towards situations which take place in Japan or in Australia.

On the other end of the scale, in the case of advanced students, early specialization occurs and will probably be retained in the future. For instance, the competence of students who graduate from an undergraduate business Japanese course will differ from that of students with an engineering Japanese background. This, too, requires a reassessment of the concept of general proficiency in the case of Japanese.

4. Spoken and written competence

Written language competence is basic in Japanese. Without the ability to read and write there can be no advanced conversation or listening comprehension. In the case of some tertiary institutions and in the case of students who study privately in Japan after having finished a limited amount of language work in Australia, there is a distinct trend to underestimate the importance of the written language.

It will be necessary for any testing instruments to pay as much attention to the written as to the spoken language.

5. Sociolinguistic competence and testing

As mentioned above, Japanese sociolinguistic competence markedly differs from Australian, British, American or any other variety of English sociolinguistic competence. Further complications result from the fact that what matters is not what these competencies are in native situations but what happens to them when they enter into a contact situation.

A considerable body of research is available on these issues and should be fully utilized for the development of testing instruments.

6. Sociocultural competence and testing

I argued above that the distance between English and Japanese norms created a need for a completely new approach to teaching and testing.

7. Assessment of presentational and performance deviation

Because of the differences in sociolinguistic and socioeconomic competence, the issue of successful communication is more difficult not only in the area of propositional communication (i.e. communication of a message), but also in the presentation of self and the amount of energy used for communication (the problem of performance).

Tests usually only deal with propositional inadequacies. However, to be able to communicate a friendly, relaxed attitude is sometimes more important for successful interaction than to convey exactly a message. Systematic inclusion of adequate attention to presentational and performance inadequacy in testing instruments is one of the important task of further development of Japanese testing.

8. Assessing learning competence

In view of the distance between the learners' and the target system a satisfactory degree of acquisition cannot normally be achieved in a Japanese undergraduate course, pass or honours. The development of independent learning competence is therefore a most important component of the overall competence of students.

This fact should be duly reflected in the testing instruments for Japanese.

CONCLUSIONS

Starting from a theory of language management, I have outlined two major avenues in which the development of Japanese language testing should proceed. I have also discussed the important concept of backwash with direct application to the Japanese situation. The paper has also listed some special Japanese problems that must be taken into consideration.

Japanese language testing in Australia lags behind the development of course design, and behind our achievements making students competent in speaking the language. It needs attention. I am glad to see at this meeting the large number of colleagues who are specialists in ESL testing and testing in other languages. Japanese language specialists will profit from their advice and I am personally looking forward to such cooperation. However, the job to develop Japanese proficiency testing is a job for people in the Japanese language teaching profession. Nothing worthwhile can be achieved by simply translating instruments developed for English or other European languages into Japanese or by fitting fragments of information on Japanese interactive competence, provided by native speakers, on previously established ESL frameworks.

Along with the development of proficiency measures it will be necessary to raise the general awareness of contemporary methods in testing throughout Australia. Perhaps this could be done in the form of a small group of experts who would present workshops on Japanese language testing in the capital cities.

(I wish to thank Helen Marriott and Robyn Spence-Brown for useful discussions of individual points included in this paper).

REFERENCES

BEFU, H. 1989. The Emic-Etic distinction and its significance for Japanese studies. Constructs for Understanding Japan, ed. by Y.Sugimoto and R. Mouer, London: Kegan Paul International, pp.323-343.

COBARRUBIAS, J. and J.A. FISHMAN, eds. 1983. Progress in Language Planning. Berlin: Mouton.

HAUGEN, E. 1966. Linguistics and language planning. Sociolinguistics, ed. by W. Bright, The Hague: Mouton, pp.50-67.

HYMES, D. 1962. The Ethnography of Speaking. Reprinted in Readings in the Sociology of Language, ed. by J.A. Fishman, The Hague: Mouton 1968, pp.99-138.

HYMES, D. 1972. Models of the Interaction of Language and Social Life. Directions in Sociolinguistics, ed. by J.J. Gumpers and D.Hymes, New York: Holt, Rinehart and Winston, pp.35-71.

INGRAM, D.E. 1985. Assessing proficiency: an overview on some aspects of testing. Modelling and Assessing Second Language Acquisition, ed. by K. Hyltenstam and M. Pienemann, Multilingual Matters, pp. 215-281.

JERNUDD, B.H. 1982. Language planning as a focus for language correction. Language Planning Newsletter 8/4, pp.1-3.

JERNUDD, B.H. 1983. Evaluation of Language Planning - What has the Last Decade Accomplished. In Cobarrubias and Fishman 1972, pp.345-378.

JERNUDD, B.H. and J.V. NEUSTUPNY 1987. Language Planning - For Whom? Proceedings of the International Colloquium on Language Planning. Ottawa: Les presses de l,Universite Laval, pp.71-84.

JERNUDD, B.H. and B. THUAN 1980. To err is human: control of language through correction. Mimeographed.

LAPKIN, S. 1985. Pedagogical implications of direct second language testing: a Canadian example. Modelling and Assessing Second Language Acquisition, ed. by K. Hyltenstam and M. Pienemann, Multilingual Matters, pp.

MARRIOTT, H.E. 1984. English Discourse of Japanese Women in Melbourne. Melbourne: Japanese Studies Centre. (Papers of the JSC 12)

MARRIOTT, H.E. 1988. Japanese Business and Social Etiquette. Melbourne: Japanese Studies Centre. (Working Papers of the JSC 13.)

MARRIOTT, H.E. 1989. Developing Australian-Japanese Tourism Research. Melbourne: Japanese Studies Centre. (Working Papers of the JSC 14).

MARRIOTT, H.E. forthcoming. Intercultural business negotiations: the problem of norm discrepancy. Journal of the Applied Linguistics Association of Australia.

MARRIOTT, H.E. forthcoming 2. The development of Japanese business communication research. Newsletter of the Japanese Studies Association of Australia 10/1, 1990.

NEUSTUPNY, J.V. 1977. Language planning for Australia. Language Sciences 45, pp.28-35. (Reprinted in Neustupny 1989.)

NEUSTUPNY, J.V. 1978. Post-Structural Approaches to Language. Tokyo: The University of Tokyo Press.

NEUSTUPNY, J.V. 1982. Gaikokujin to no komyunikeeshon. Tokyo: Iwanami Shoten.

NEUSTUPNY, J.V.. 1983. Towards a paradigm for language planning. Language Planning Newsleter 9/4, pp.1-4.

NEUSTUPNY, J.V. 1985a. Problems in Australian-Japanese contact situations. Cross-Cultural Encounters, ed. by J.B. Pride, Melbourne: River Seine, pp.44-64.

NEUSTUPNY, J.V. 1985b. Language norms in Australian-Japanese contact situations. Australia, Meeting Place of Languages, ed. by M. Clyne, Canberra: Pacific Linguistics, pp.161-171.

NEUSTUPNY, J.v. 1987. Communicating with the Japanese. Tokyo: The Japan Times.

NEUSTUPNY, J.V. 1988a. Australia and Japan: Cross-cultural communication problems. Vox 1, pp.28-32.

NEUSTUPNY, J.V. 1988b. Problems of English contact discourse and language planning. Paper presented at the 1988 Regional Seminar on Language Planning in a Multilingual Setting, National University of Singapore.

NEUSTUPNY, J.V. 1989. Strategies for Asia and Japan Literacy. Melbourne: Japanese Studies Centre (Papers of the JSC 15).

NEUSTUPNY, J.V. and A. OZAKI 1986. Intaaakushon no tame no Nihongo kyooiku. Nihongo kyooiku 59, pp.117-134.

RUBIN, J. 1978-9. The approach to language planning within the United States. Language Planning Newsletter 4/4, pp.1-6 and 5/1, pp.1-6.

RUBIN, J. and B.H. JERNUDD 1971. Can Language be Planned? Honolulu: The University Press of Hawaii.

THOMPSON, I., R.T. THOMPSON & D. HIPLE 1988. Issues concerning the less commonly taught languages. Second Language Proficiency Assessment, ed. by P. Lowe and C.W. Stansfield, Prentice Hall, pp. 83-123.

VACHEK, J. 1966. The Linguistic School of Prague. Bloomington: Indiana University Press.

VALLETTE, R. 1977. The culture test. Reprinted in Culture Bound, ed. by J.M. Valdes, C.U.P. 1986, pp.179-197.

DESIGNING INSTRUMENTS TO MEASURE LANGUAGE PROFICIENCY

ALAN DAVIES and ANNIE BROWN

University of Melbourne

We have been asked in this International Working Group to begin our consideration on the assessment of Proficiency Levels where all such discussion should start, that is with the question of purpose. That is entirely proper. All test construction and development should begin with the question of the purpose for which the test is needed. Furthermore, and again quite properly, we have been offered a model (which is helpfully variable) of proficiency in the shape of sets of proficiency rating scales. I do not question the importance of beginning with the issue of purpose nor with the proficiency model presented to us. The question I do raise is the "how?" question. By what methods do we achieve our purpose of establishing proficiency? Testing enterprises require that purpose and model be settled but neither is a substitute for a methodology of testing. Neither is a test. What we still need to do is to examine the variety of methods available to us for establishing proficiency. That is the basic question I address in this paper.

In so doing I will make three points relating to:

(i) the use of Proficiency Scales in language testing;

(ii) the need to take full advantage of all the methods available for providing assessment information;

(iii) the importance of work on construct validity of proficiency tests of Japanese as a Foreign Language, work which we have started making a contribution to in the Language Teaching Unit at the University of Melbourne. But before these three points, I shall make a brief preliminary observation.

In a recent radio interview Sir Claus Moser, current President of the British Association for the Advancement of Science, remarked that after a decade or more of rejection by governments in many countries social science is now making a comeback and is being regarded as having something of importance to say about future policies. He noted that this change in regard is evident among politicians and among natural scientists (witness his own presidency of the BAAS, the first holder of that office to be a social scientist). He commented that in part this is because so many of the problems we face are social, giving as an example the present world-wide concern with global warming (whether misconceived or not). Sir Claus is himself a distinguished demographer and it is unsurprising perhaps that he would reflect in this way, given that so many of our future immediate problems arise from the increase in the world population (which has doubled in the last 30 years).

But for our purposes today, if it is indeed the case that social science (broadly interpreted) is to be seen as of more relevance to social and national concerns, all the more reason for making as sure as we can that the information, the research, the results we provide for others to act upon are as accurate as possible, that they reflect the complexity both of human beings and of social settings - a point Sir Claus also referred to, mentioning that social science is much more difficult to be precise about than is natural science, simply because people are far more varied, far less systematic than is the universe.

1. The increasing use of proficiency scales in language assessment has both positive and negative aspects. On the positive side they are authentic examples of language in use; there is no gap between what Bachman (1990) calls 'the criterion of proficiency and the definition of authenticity' (1990:409). Because such tests are typically direct authenticity comes as it were with the test and does not have to be appealed to or claimed elsewhere. It is therefore often argued on behalf

of such techniques that they have built-in validity. On the negative side it must surely be pointed out that all tests (not just indirect or semi-direct ones) lack authenticity. They are all simulations of real life rather than real life itself. What this of course suggests is that it is the job of assessment not so much to replicate real life (because by definition that cannot be done or when done is always partial and potentially biased) but to reflect language learning abilities and to sample real life situations rather than to collect them The old example from the testing of reading makes the point forcibly it is surely clear that when we test someone's literacy on a text we have no serious interest in that particular text. What we are interested in is the learner's ability to read texts. How important therefore it is that the text used for the test, and the tasks required in the test should be adequate samples of texts at the appropriate reading level and of the tasks required in that reading.

Tests which make use of proficiency scales frequently use the interview as a means of sampling language data which can then be related to the scale. Interviews are said - as I have hinted above - to provide direct entry into the speaker's language ability. But they are also notoriously unreliable. Yes, there are ways of training judges and ways of pooling judgements but the extent to which they can be made less unreliable is a function of the amount of training and time and money that are available. In other words *plus* reliable means *minus* practical. As Ingram (this conference) points out these are issues of validity and of reliability and to some extent overriding both is the issue of practicality. Which is to say that the nearer the test approximates to real life (the longer it is, the more assessors there are and so on) the more validity and the more reliability it will have. But then it ceases to be a test and becomes precisely what it is supposed to be predicting, which is real life use.

Direct tests are about truth not really about testing. They confuse the criterion with the test. These must never be confused since the test necessarily is an icon of the criterion and not the criterion itself. Directness in language testing, like authenticity, represents an extreme flight from behaviourism, but also like behaviourism directness is just as extreme because at the opposite pole. A compromise is needed.

Bachman (1990) describes such a compromise in his useful discussion of face validity and of direct tests (in his view direct tests are basically attempts to embody face validity, a concept he dismisses as not serious either academically or pedagogically). If face validity (and therefore direct tests) have any respectability it can only be in single settings. What that means is that a new direct test must be administered for every new setting, every new situation which they are meant to reflect. They have no generalisability, which Bachman argues must be central to all testing. The purpose of a test is to provide a sample of authentic language behaviours, not, note, language use or even language behaviour. In other words a test must provide information about the abilities that underlie language performance in real life situations not that performance itself. Bachman offers the following way forward:

'rather than attempting to replicate actual uses of non-test language' (that is real-life) 'use, the interactional/ability (IA) approach uses a theoretical framework of factors that affect test performance to construct tests that include features of language use that are relevant to the interpretations and uses that are to be made of test scores. Our major interest thus shifts from that of attempting to mirror non-test language

use to that of determining what combination of test method facets is most likely to promote an appropriate interaction of a particular group of test takers with the testing context. The second difference is that unlike the Real Life (RL) approach the IA approach must necessarily consider the abilities of the test taker. That is, in the RL approach language ability need not be a consideration, since it is performance on real-life tasks that provides the de facto definition of ability' (1990:430/1).

2. The useful definitions of proficiency provided by Proficiency Rating Scales, such as those we have under consideration, do not in themselves provide any principled reason for avoiding or demoting the necessary concern in language testing with test method (or 'test method facets' as Bachman calls them). Indeed one of the major, and yet in principle irrelevant, objections to the wholesale use of Proficiency Scales in oral assessment is that the method typically used is that of the interview. Both the construct and the definitions of proficiency still have just as much relevance and value when we make use of other methods of providing ourselves with the necessary information about the learner's language skills.

There is in direct tests such as the interview the ever-present danger of routinisation (as indeed there is in the analogue, communicative language teaching). A good example of the danger can be found in the so-called *neck verse*. In the Middle Ages in England (and perhaps elsewhere) the Church had a separate legal jurisdiction from the state. Its clergy were not subject to the jurisdiction of the state. Since the Church had less severe punishment (notably there was no capital punishment, at a time when very many offences were punished by death by the state) it was obviously to a person's advantage to be tried for an offence by the Church by claiming benefit of clergy. This required summary proof of being a cleric. Since that was not an easy matter (how did you prove you were a cleric before the days of printing and telephones, of anger and of telegrams?) appellants were required to prove their clerisy to the state officials by reading a verse from the Bible. That was a direct test and no mistake. However, since they often had to read (in Latin of course) to people who were themselves either illiterate or only partially literate the practice grew up of always being required to read the same verse, (from Psalm 51), the so-called neck verse, the one that begins: *'Miserere mei, deus'* ('Have mercy upon me, O God, after thy great goodness: according to the multitude of thy mercies, do away my offences'). Such routinisation led, as is obvious, to the abuse of many who were not clerics claiming the right to be tried by the Church for an offence. But that is the danger inherent in all direct tests, that in order to be fair, to avoid subjectivity, the test becomes more and more routine as time goes on and eventually as little like real life as the most indirect test but without its special claim to be a sample of underlying language skills.

Furthermore, interviews in practice depend on assessors' impressions which themselves require the combination of discrete items. True they are supposed to allow for spontaneity, but in view of my example of the neck-verse do they always do so?

As far as the proficiency scales themselves are concerned we cannot avoid the basic question which is just how valid they are. There is a sense in which all such scales are circular. The fact that they bring together proficiency level with authenticity is assumed to be itself an indication of value - but the question remains of just how valid the levels are. In the physical world we can indeed divide up nature in equal units and claim

that the units are recursive (for example in measuring height) that each one is the same as the next. But just what is the status of the descriptors and of the example tasks at each level of the Proficiency Scales on offer. Are they predictors or are they criteria? How valid are they?

To move from the Church and the law to that other great profession, medicine, consider the relation which we accept, most of us, in medicine between the doctor's diagnosis and the health we wish to have. Health (like proficiency) can be put on a scale so that we are all more or less healthy. However, we do not welcome or indeed expect the doctor who examines us (whether when we are unwell or when we need an insurance policy) to use only direct tests of health, for example getting us to run on the spot. What we want our doctor to do is to obtain in as precise a way as possible the best information about our body functions which s/he can relate to a health scale. In other words we are indifferent to the directness or indirectness of the medical tests used by a doctor to determine our body health. Why should an investigation into our language health be any different?

We must remind ourselves of just what a test is. A test has three requirements. It must:

(a) sample the domain it lays claim to;
(b) discriminate in the sense of distinguishing among individual abilities.

It is for these reasons that we distinguish between tests and teaching, a distinction that face validity does not always make. A test is not an exercise on which to practise language learning. As we have already seen the sampling of a test relates to its validity; the discrimination to its reliability.There is also the third requirement which we mentioned earlier. A test must:

(c) be practical and not require time or resources that are unobtainable or impossible.

What a test does is to predict a criterion. In my view the Proficiency Scales we have under discussion are a type of criterion. The interview is one such test method facet. There are many others and it would be wasteful and lacking in principle to the extent of over-enthusiasm if we were to restrict ourselves to a one type method. (In politics the criticism has been made of Treasury Ministers who concentrate on one method of fiscal policy that they are playing a 'one-club game'!) The testing literature provides a wide range of test methods and techniques. Construction of a set of proficiency measures for JFL, and in due course for other Asian languages, should make full use of the array of test types and methods now available.

As an example of an array of test methods or test method facets which were designed with purposes similar to our own today I would like to refer to the UK ARELS tests (or to give them their current full title 'The Oxford-ARELS Examinations in English as a Foreign Language'). These examinations represent a coming together of two sets of tests, the ARELS tests of Spoken English which have been available now for over 15 years and the Oxford Delegacy tests of non-literary writing and reading skills based on authentic tasks. I will refer solely in what follows to the tests of the spoken language. (Quotations are from 'The Oxford-ARELS Examinations in English as a Foreign Language, Rationale, Regulations and Syllabuses', Oxford 1989: 5-7

'The particular thrust of the (tests) is towards the improvement of skills

in the use and comprehension of spoken English. These are notoriously difficult to measure objectively, reliably and validly and ARELS Oral Examinations employ techniques which show marked improvement on conventional methods of oral examining. A further characteristic of these examinations is the importance attached to the use of English in everyday, realistic situations. At all levels Spoken English is examined as a means of communication rather than as an academic exercise.'

'The examinations are set at three levels: Preliminary, Higher and Diploma. A Junior version of the Preliminary level is also available.

The examinations are held in language laboratories. The candidate listens to a Master Tape (the examination) and responses are recorded on a personal tape - the equivalent of an answer paper in a written examination.

The examination is identical wherever in the world it is held because the Master Tape is unchanged.

The sample of Spoken English obtained from each candidate is extensive. In the Diploma (the most advanced of the three levels) it is about 30 minutes long.'

(The booklet gives details of the moderation meetings, of marking (double marking at the two higher levels), and of the detailed Marking Keys which are made available for each diet.)

'Because the examinations assess performance rather than factual knowledge it is not necessary for all candidates to take the examination at the same time or on the same date'

Full details of the Performance Objectives for the Preliminary Level are given in Appendix 2 to this paper. Also in Appendix 2 are examples of the questions and of the Marking Guide set for the 1989 Test AP31 (the 'Preliminary Examination in Spoken English and Comprehension.')

3. The papers circulated prior to this meeting did indeed refer to other than direct tests, semi-direct and indirect approaches. In the LTU we have considered this issue with particular reference to the testing of Japanese as a FL in Australia in skills relevant to a variety of tasks among workers in tourism and related industries. I restrict myself on this occasion to the consideration of a test method which approximates most closely to the direct test type we have under discussion. What matters after all is that the test should be carefully designed and structured so as to sample the skills and abilities relevant to the domain under test. In other words, proficiency tests need construct validity: content validity alone is insufficient.

3.1 Rationale for the test

The growth over the last few years in the number of overseas visitors to Australia, particularly from Asian countries, and the projected continued growth, has meant that the demand for people with language skills to work in the area of tourism will continue to rise.

Whilst training courses in languages relevant to the needs of the tourist industry (Japanese in particular, but also Korean and Chinese) are proliferating at all levels of the education system, there is at present no foreign language proficiency assessment scheme specifically catering to the

needs of the travel and tourism industry within Australia. If language qualifications are to be prerequisites to gaining employment, then both the curricula of the language courses offered and the form of proficiency assessment must reflect the needs of the industry[1].

To this end the Language Testing Unit of the National Languages Institute of Australia proposes the development of testing tools linked to proficiency rating scales for a range of Asian languages. This test battery will reflect the needs of the tourist industry in that it assesses the language skills relevant to the work situation, and at the same time presents future employers with clear information on the candidate's language ability in terms of these work-related skills.

A National Language Assessment Scheme is seen to be of vital importance for two reasons. Firstly, it will allow standards of language training to be applied across a range of occupations within the tourist industry throughout Australia, a necessity in a dynamic and developing employment area. Secondly, it is hoped that such a scheme will have some washback effect on the syllabuses of language training courses, thus producing graduates with language skills appropriate to the tourist industry, giving employers access to a wider range of prospective employees, and enabling them to employ young Australians in areas where previously they have had to employ foreign nationals.

3.2 Test Users

It is envisaged that the test will be applicable within the tourist industry in Australia (and to some extent within commerce). The need for nationally recognised proficiency standards to be established across the whole of the industry has prompted the design of a test which is not narrowly occupation specific but relates to the industry as a whole.

The test may be used to screen for employment where language skills are required, it may be used to make decisions on further language training, or it may be used as a means of assessing for recognition of language skills.

There will be two levels of the test - Basic Skills and Advanced Skills. The Basic Skills level assesses in the range ASLPR 0 - 2, and the Advanced Skills level assesses in the range ASLPR 2 - 5.

Occupational areas for which the test is appropriate include customer contact areas such as:

Airline:	Phone Reservations
	Cabin crew
	Airport Check-in Staff
	Travel Centre
Travel:	Agency staff
	Tour guides
Hotel:	Reception
	Service - sales, room service, food and drink, recreation
	Business facilities
Banking	Foreign exchange
Education	Administration, welfare
Sales	Duty Free shops

---

[1] Kinnaird B. Paper presented at National Assessment Consultation, Sydney, December 1989

The Advanced Skills level also assesses candidates' level of proficiency in language skills appropriate to the occupational requirements of management and commerce.

3.3 Test Format

The most appropriate format for the test is one which reflects the types of communication skills required of the candidates in providing service to clients within their occupations. Interpersonal language skills and information exchange skills are of prime importance at the Basic Skills level.

Thus the test will consist of a series of tasks which enable the candidates' proficiency in these areas to be assessed.

Four formats for the oracy test have been considered.

1. A direct test (live interlocutor), where the assessor observes the interaction between candidate and interlocutor.

2. A direct test where the interlocutor is the assessor.

3. A direct test where the interaction is taped and assessed centrally.

4. A semi-direct test, where there is no live interlocutor, the stimulus being provided on tape or in textual form, which is assessed centrally.

The advantages and disadvantages of each of these will be considered below.

3.3.1 Practicality

3.3.1.1 Test availability

A centralised marking scheme enables the test to be administered on demand anywhere in Australia. This would be impossible where on-site recruitment, selection and training of assessors were a necessity.

3.3.1.2 Training of assessors

A centralised marking scheme allows for a small number of assessors to be trained. Given the specialised type of person required to carry out this role - native-speaker of the language, language teacher-trained and accredited as an assessor, the problems involved in selecting and training such people on a nationwide basis, ensuring that they maintain their skills through regular test involvement, and finding and training replacements as individuals become unavailable, lead us to rule out options 1 and 2 as being impractical and restrictive. This would also apply to some extent to option 3, the interlocutor necessarily being a native-speaker of the language and with some awareness of elicitation and interaction strategies.

3.3.1.3 Cost-effectiveness

A semi-direct test can be administered in a language laboratory setting to several people simultaneously. A direct test, however, must be administered individually, and will therefore be more costly.

Centralised marking allows a small pool of assessors who work on a regular basis to be maintained, thus minimising the costs of training and re-training.

Assessment time for a direct test is at least the length of the interview, which includes both candidate and interlocutor talk time. In an indirect test one can ensure that only the candidate's responses are recorded onto the assessment tape, which cuts down the listening time for the assessor, and hence the overall time for assessment.

3.3.2 Validity

A test designed to assess oral proficiency should reflect the types of speech situation to be encountered in real-life use. A direct test obviously has a higher degree of face-validity than a semi-direct one. However, even in an interview situation the communication does not reflect real life - the participants are fully aware that they are examiner and candidate and not, say, receptionist and guest. Furthermore, given the wide range of short communication tasks proposed for this test as being typical of the types of occupational interaction, it is felt that in a one-to-one interaction role-switching may seem more unrealistic than on a tape-based test where a range of stimulus voices and acoustic conditions can be provided, thus effectively separating for the candidates each task and role, and enabling them to feel more comfortable with the tasks.

3.3.3 Reliability

In a tape-based test each candidate has the same stimulus and thus the same opportunities for demonstrating his/her capabilities, in contrast to direct tests of oral proficiency where the problem of inter and intra-eliciter reliability cannot be avoided.

A tape-based test should have scoring reliability higher than a direct test, the scoring procedures consisting of a mixture of atomistic rating scales as well as objective scoring, due to the range of task types available.

There are problems inherent in a direct test where a range of short tasks are to be performed. Firstly, the examiner has to organise, initiate and explain each task. It is inevitable that this will be a slower process than in a semi-direct test. Furthermore, given the complexity of the organisational requirements and the examiner's involvement in the role-plays, it would be unwise to assume that she can effectively constantly monitor the candidate's output. To ensure reliable assessment one would require an interlocutor and an assessor, which would obviously make the test extremely expensive to administer, or one would need to record the interaction for later assessment. In a semi-direct test this problem does, of course, not arise.

3.4 Conclusion

Practical constraints lead us to rule out options 1 and 2 above. Options 3 and 4 will be investigated in more detail in terms of validity and acceptability to users.

Appendix 1

Tape-based tests

It should be noted that a semi-direct tape-based test format has been used successfully for several years in such tests as the ARELS Examinations in Spoken English and Comprehension, the Test of Spoken English (TSE) (Clark 1979), the Test in English for Educational Purposes (TEEP) (James 1988), the Portuguese Speaking Test PST) (Stansfield et al 1990 and Cohen 1989) and the Recorded Oral Proficiency Examination (ROPE) (Lowe and Clifford 1980).

Appendix 2

[For copyright reasons the material has not been included in this publication.]

Bibliography

ARELS Examinations Trust (1989) The Oxford-ARELS Examinations in English as a Foreign Language: Rationales, Regulations and Syllabuses Oxford ARELS Examinations Trust, Ewart House, Summertown, Oxford OX2 7BZ, U.K.

Bachman L.F. (1990) Fundamental Considerations in Language Testing OUP Oxford

Clark, J.L.D. (1979). Direct vs. Semi-Direct Tests of Speaking Ability. In E.J. Briere and F.B.Hinofotis (eds), Concepts in Language Testing: Some Recent Studies. Washington D.C.: TESOL

Cohen, A.D. (1989) Portuguese Speaking Test: Test Review. In Language Testing Vol 6,2. Bristol: Edward Arnold.

James, G (1988). Development of an Oral Proficiency Component in a Test in English for Academic Purposes. In A. Hughes (ed), Testing English for University Study. ELT Documents: 127. Oxford: Modern English Publications and the British Council.

Lowe, P and Clifford, R.T. (1980) Developing an Indirect Measure of Overall Oral Proficiency. In J.R. Frith (ed), Measuring Spoken Language Proficiency. Washington D.C.: Georgetown University Press, 1980. [ROPE TEST]

Stansfield, C.W. et al (1990) The Development and Validation of the Portuguese Speaking Test. In Hispania (forthcoming Sep/Dec 1990).

DAILY TASKS AND OUTCOMES

OUTCOME FROM DAY 1

Needs Identified

Industry needs

Comprehensible global proficiency scales: general
special purpose

Measures of employee skills for award purposes, salary loadings, productivity payments

Comprehensible general proficiency assessments for task allocation, training and re-training purposes

Highly task specific assessments at 0 to 1- for people operating with largely formulaic utterances in strictly defined and limited range of predictable situations

Measures of ability to operate in specific activities requiring more creativity (e.g. secretarial, tour guides, etc.): general proficiency in defined contexts.

Self-assessment procedures

Incentives for further learning

Information on duration of language courses to attain needed proficiency levels

Wide range of fields to be catered for, including:

science, technology, office work, hotel reception, tour guides for commentary and liaison on services, marketing, administration, business, law, public service, teaching, tourism, hospitality, engineers, architecture, trade, commerce, real estate, retailers, shop assistants, finance, sub-titlers, film industry personnel, foreign postings and other relevant activities for Departments of Foreign Affairs, Trade (including Austrade), Immigration, Defence and other relevant departments. NB: Potential data from Immigration records; need to be able to group and rationalise these areas, cf. IELTS.

Assessment of language of Australians working for Japanese companies in Australia or Japan

Promotion of realisation of the need for tests

Assessment of cultural understanding, cross cultural attitudes irrespective of language skill or participation in language program

Appropriate cultural attitudes

Procedures to identify the proficiency needed for a wide range of tasks, the proficiency of the people to undertake them: continuum of needs, continuum of levels for individuals

Separate focus on speaking, listening, writing and reading in response to different needs of industry

Proficiency testing as a screen for other more specialised language activities (e g. before taking interpreter/translator courses or tests)

Aptitude assessment to screen employees prior to language training

Quality control/consumer protection on courses that employees may take

School needs

Proficiency scales: general
special purpose

Measures of student progress in courses

Means of diagnosing student strengths and weaknesses

Measures of whether students exiting from schools can operate in certain specific areas as well as with some degree of general proficiency.

Self-assessment procedures

Incentives for further learning

Information on duration of language courses to attain needed proficiency levels

Means of knowing what proficiency students have irrespective of where they have studied: need for a national assessment available to or required for all school graduates

Quality control on teachers

Proficiency framework within which courses and series of courses can be coherently designed

Activities related to learner proficiency

Curricula that acknowledge/are based on developmental continua as in proficiency scales

Separate focus on speaking, listening, reading and writing

Goals, programs related to general proficiency plus possibly specific purpose areas (e.g. developed in immersion programs, self-access, work experience, TAFE, etc.)

Means of assessing students on entry (streaming/placement), through a course and on exit

Aptitude assessment

Certification

Quality control on courses

Tertiary needs (TAFE/Higher Education)

Proficiency scales: general
special purpose

Incentives for further learning

Information on duration of language courses to attain needed proficiency levels

Means of knowing what proficiency students have irrespective of where they have studied: need for a national assessment available to or required for all graduates

Self-assessment procedures

Measures of ability to cope with life abroad and academic study in Japanese medium courses

Identification of amount of pre-sessional Japanese language study required to cope with demands of tertiary courses

Streaming/placement

Activities related to learner proficiency

Curricula that acknowledge/are based on developmental continua as in proficiency scales

National (international) system of proficiency scales, tests, etc.

Aptitude assessment

Certification

Quality control on teaching and courses

Proficiency scales, tests, etc. that take account of sociocultural phenomena.

General issues

Certification

Duration of validity of proficiency test result

User driven and user friendly scales, tests and results

Implications of student ages

Appropriateness of test types to purpose and context

TASKS FOR DAY 2

1. Consider and rationalise needs identified on Day 1.

2. Identify the scales, tests, etc. required to meet the needs determined. (David Ingram's paper lists test types, etc.)

3. Start to sketch specifications for the scales and tests proposed. Some issues to be considered in specifications include:

   purpose of the test
   candidates
   focus area of the language
   macroskill (S,L,W,R)
   focus proficiency level
   topic range
   specific tasks for focus
   cultural issues
   item types
   context of administration
   who, how, when, where to administer

   (See paper on IELTS in portfolio for IELTS specifications, p. 10)

4. Identify order of priority for test instrument development.

OUTCOME FROM DAY 2

Instrument Needs

Priorities

1. Generic scale: Japanese (eventually also Chinese, Indonesian, Korean and Thai)
2. Speaking instrument (related to generic scale)
3. Listening (related to generic scale)
4. Reading (related to generic scale)
5. Writing (related to generic scale)
6. Role of grammar in proficiency assessment (related to generic scale)
7. Japanese for inbound tourism - multiple instruments at two levels

   (a) 0 to 1 and
   (b) 1+ to 3

8. Japanese for business/law (instruments to relate to special purpose scale in range 1+ to 3)
9. Japanese for science and technology (instruments to relate to special purpose scale in range 1+ to 3)
10. Version of the scale that can be used for special purpose assessment
11. Proficiency assessment in the school context
12. Diagnostic tests (related to national curriculum)
13. Attainment tests (related to national curriculum)
14. Japanese for academic purposes:

    -adding speaking and writing sub-tests to the Japanese Language Proficiency Test

15. Assessment of sociocultural and sociolinguistic knowledge
16. Self assessment procedures
17. Certification of learners
18. Certification of teachers
19. Aptitude assessment

TASKS FOR DAY 3

The aim is to develop briefs for specific test instrument development projects including such issues as:

1 management of the project

2. how the project is to be undertaken

3. stages of the project including, for example:

   - needs analysis and related research
   - client group
   - item types
   - (possibly indication of likely test content)
   - validation procedures

4. possible applications

5. time-lines

6. how and by whom the test is to be administered

7. scope for international cooperation

8. costing

Each group is asked to develop project profiles for the test instruments in the order listed. In some cases the project profiles would be similar.

GROUP 1

Speaking instrument (related to generic scale)

Listening (related to generic scale)

Reading (related to generic scale)

Writing (related to generic scale)

Role of grammar in proficiency assessment (related to generic scale)

GROUP 2

Japanese for academic purposes

   - adding speaking and writing sub-tests to the Japanese Language Proficiency Test

Assessment of sociocultural and sociolinguistic knowledge

Aptitude assessment

GROUP 3

Proficiency assessment in the school context

Diagnostic tests (related to national curriculum)

Attainment tests (related to national curriculum)

GROUP 4

Generic scale

Version of the scale that can be used for special purpose assessment

Japanese for inbound tourism - multiple instruments at two levels

(a) 0 to 1 and

(b) 1+ to 3

Japanese for business/law (instruments to relate to special purpose scale in range 1+ to 3)

Japanese for science and technology (instruments to relate to special purpose scale in range 1+ to 3)

OUTCOME FROM DAY 3

See next section RESULTANT PROJECT PROPOSALS

RESULTANT PROJECT PROPOSALS

Note: The proposals in this section relate to Instruments 1 - 15. The remaining four instruments identified were considered less likely to attract funding in the foreseeable future. However, projects to meet these needs could be considered (and proposals such as those in this section developed) for Higher Degree research.

PROJECT PROPOSAL 1

DEVELOPMENT OF THE GENERIC SCALE
(relating to Instrument 1)

Aim:

To develop a generic scale as a basis for the assessment of proficiency in Japanese and as an overarching framework within which curricula may be rationally and coherently developed.

Since a great deal of work has already been devoted to the development of proficiency scales in Australia (i.e. the Australian Second Language Proficiency Ratings) and elsewhere (e.g. the ACTFL scale in the United States, which in fact drew substantially on the ASLPR), it is recommended that this project finalise the development of the Japanese version of the ASLPR. This proposal has the advantage of being most economical and of facilitating comparison between the new scale, the ASLPR, the ACTFL scale and other existing scales.

It is proposed to call the new scale the Asian Languages Proficiency Rating Scale (Japanese).

The new scale will serve as a model for parallel scales in other Asian languages.

Highest priority will be given to development of speaking and listening sub-scales though reading and writing scales would also be required.

Procedures:

Before commencing any project presently under consideration, the relationship between the National Languages Institute of Australia, the Asian Studies Council, industry and other users of the tests should be clarified.

1. Establishment by the Asian Studies Council of a small working party to draw up terms of reference for this project and to consider the relationship between the development of the generic scale and of Instruments 2 to 6 - the tests used to relate learner proficiency to the generic scale (Project Proposal 1). There may be some cost advantage in these being combined into a single project. There may also be advantage in linking this total project to development of the special purpose scales and related tests, though these instruments have a slightly lower priority.

2. Call for expressions of interest and allocation of the project.

3. Formation of a project advisory committee to include representatives of key users (including employers), experts in language testing (especially proficiency assessment), native speakers of Japanese, and highly proficient native English speaking teachers of Japanese.

4. Survey and review of other scales, relevant instruments and relevant literature (including developmental studies of Japanese as a second or foreign language).

5. Review of the existing version of the ASLPR for Japanese and its initial trialling in face-to-face interviews with approximately 100 learners of Japanese throughout the proficiency range in different courses, in different locations, and of different ages from adolescents up.

6. Formal trialling and validation using the test procedures being established in the development of Instruments 2 to 6.

7. Review and modification of the scale in the light of the trial results.

8. Release of the Asian Languages Proficiency Rating Scale (Japanese).

Timelines:

Given that there is close interrelationship between this project and the project to develop Instruments 2 to 6 with input from one to the other, the likely time required is estimated at 18 months with an absolute minimum of 12 months. However, if the two projects are developed as a single project with no additional staffing allocated, it would be reasonable to estimate that an additional time of 6 months over and above that indicated for the development of Instruments 2 to 6 would be required (i.e. a total time of 2 years).

Indicative staffing:

Project Director) Testing Expert)	equivalent to 0.2 of a full-time position
Project Coordinator (an applied linguist who is a native or high proficiency speaker of Japanese with an understanding of proficiency rating scales)	1 full-time position
Secretary (bilingual)	0.2 of a full-time position
Advisory Committee (see 3 under Procedures above)	

Indicative costing (estimated for 18 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent .2 of a full-time position)	$ 29,000
Project Coordinator (estimated at middle of lecturers scale)	$ 75,000
Bilingual secretary	$ 11,000

TOTAL SALARY COSTS		$ 115,000
Advisory Committee costs for 4 meetings	estimated	$ 10,000
Research costs:		
- literature reviews and incidentals	estimated	$ 2,000
- consultations with user groups and institutions and informants	estimated	$ 6,000
Development interviews estimated cost including travel and informant time	estimated	$ 6,000
Trials:		
- preparation of video and other materials	estimated	$ 5,000
- travel	estimated	$ 6,000
- computer costs: contributed by institution		-
Consumables	estimated	$ 2,000
Equipment	estimated	$ 2,000
Office space: contributed by institution		-
Institutional overheads	nominal	$ 25,000
	TOTAL	$ 179,000

Funding sources:

In addition to Australian government sources, contributions may be available from Canada, Australian industry, the Japan Foundation and Japanese industry.

PROJECT PROPOSAL 2

TESTS REQUIRED TO RELATE LEARNERS TO THE GENERIC SCALE
(relating to Instruments 2 to 6)

Stages of the project:

Development of specifications and trial version of tests (significant research required, especially for speaking and writing)
estimated time 12 months

Trials and development of interpretative procedures
estimated time 4 months

Review of trial tests and development of the first live version
estimated time 2 months

Outcomes:

One live version of the tests to be secure and centrally administered

Advice to other users on development of their own tests

Additional items to place in test item banks

Indicative staffing:

Project Director) equivalent to 1 full-time
Testing Expert) position

Test Writers (applied linguists who are native or high proficiency speakers of Japanese) 4 full-time positions

Secretary (bilingual) 1 full-time position

Advisory Committee (largely local teachers, applied linguists and user representatives plus an ASC representative)

Indicative costing (estimated for 18 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent to 1 full-time position)		$ 120,000
2 test writers (estimated at middle of lecturers scale)		$ 188,000
Bilingual secretary		$ 52,000
TOTAL SALARY COSTS		$ 360,000
Advisory Committee costs for 4 meetings	estimated	$ 10,000
Research costs:		
- literature reviews and incidentals	estimated	$ 4,000
- consultations with user groups and institutions	estimated	$ 6,000
Trials:		
- travel	estimated	$ 6,000
- computer costs: contributed by institution		-
Consumables	estimated	$ 5,000
Equipment	estimated	$ 8,000
Marketing	estimated	$ 3,000
Office space: contributed by institution		-
Institutional overheads	nominal	$ 40,000
TOTAL		$ 442,000

It is estimated that the test should become self-funding within a period of two years after release of the first live version of the test. Consequently some on-going development costs will need to be met if additional parallel versions of the test are to be available for reasons of security. The estimated cost of the on-going development activities is 50% of the development costs in the first year after release of the first live version

of the test and 25% in the second year, making a total additional cost of $331,500.

Total development costs before the test becomes fully self-funding $773,500.

Note: This project is closely related to the development of the generic scale (Project Proposal 1) and there will be substantial saving if the two development activities proceed together as a single project.

PROJECT PROPOSAL 3

TESTS OF JAPANESE IN THE AREAS OF TOURISM AND HOSPITALITY (INITIAL PROFICIENCY)

(relating to Instrument 7a)

Aim:

To develop tests of proficiency in Japanese for personnel in the areas of tourism and hospitality. The test development project outlined here will serve as a model for other tests of this sort, though rationalisation across areas may lead to some economy.

Test Focus:

These tests are intended for use with low level learners in the range of ASLPR S:0, L:0 to S:1, L:1 whose language consists largely of rote-memorized formulae for use in defined situations. Reading and writing are excluded, though subsequently reading tests could be included if required.

Stages of the project:

Review of literature on Japanese for tourism and hospitality and of any courses in these areas, and consultations with industry to establish highest priority areas

estimated time one month

Needs analysis in designated area, i.e. highest priority area identified by consultations

estimated time two months

Development of specifications, assessment criteria and trial version of tests (S and L). S to be assessed either by live interview or SOPI in parallel forms

estimated time three months

Trials and development of interpretative procedures

estimated time one month

Review of trial tests, development of first live version, and development of interviewer training package

estimated time one month

Outcomes:

One live version of the tests to be secure and centrally administered

A model allowing economical and efficient production of subsequent tests of a similar type in the same proficiency range in the area of tourism and hospitality.

Indicative staffing:

Project Director) Testing Expert)	equivalent to 0.2 of a full-time position
Test Writers (applied linguists who are native or high proficiency speakers of Japanese)	2 full-time positions
Secretary (bilingual)	0.2 of a full-time position
Advisory Committee (largely native Japanese speaking teachers, applied linguists and teachers of Japanese plus an ASC representative)	

Indicative costing (estimated for 8 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent to 0.2 of a full-time position)		$ 12,000
2 test writers (estimated at middle of lecturers scale)		$ 65,000
Bilingual secretary		$ 5,000
TOTAL SALARY COSTS		$ 82,000
Advisory Committee costs for 2 meetings	estimated	$ 4,000
Research costs:		
- review of literature and courses and incidentals	estimated	$ 2,000
- consultations with user groups	estimated	$ 2,000
- needs analysis	estimated	$ 5,000

Development of interviewer training package	estimated	$	6,000
Trials:			
- travel	estimated	$	6,000
- computer costs: contributed by institution			-
Consumables	estimated	$	3,000
Equipment	estimated	$	5,000
Marketing (N.B. No salary costs included)	estimated	$	5,000
Office space: contributed by institution			-
Institutional overheads	nominal	$	16,000
TOTAL		$	136,000

It should be relatively easy for a centralised administration body to maintain security on this test. It is also estimated that the test should become self-funding soon after release of the first live version of the test. Consequently no on-going development costs are included for additional parallel versions.

This test may be marketable in other countries that receive Japanese tourists (e.g. South East Asia, Europe and North America) and this may help to recover initial and on-going development costs.

Note: This project is closely related to the development of the test in the 1+ to 3 range (Project Proposal 4) and there may be some saving if the two development activities proceed together as a single project.

PROJECT PROPOSAL 4

TESTS OF JAPANESE IN THE AREAS OF TOURISM AND HOSPITALITY. (INTERMEDIATE PROFICIENCY)

(relating to Instrument 7b)

Aim:

To develop tests of proficiency in Japanese for personnel in the area of tourism and hospitality. The test development project outlined here will serve as a model for other tests of this sort, though rationalisation across areas may lead to some economy.

Test Focus:

These tests are intended for use with learners in the range of ASLPR S:1+, L:1+ to S:3, L:3 in specified situations where, nevertheless, creative use of the language (rather than the use of rote-memorised formulae) is required. Reading and writing are excluded, though subsequently reading tests could be included if required. Speaking assessment will be made using the Asian Language Proficiency Ratings Scale.

Stages of the project:

Review of literature on Japanese Tourism and Hospitality and of any courses in these areas, and consultations with industry to establish highest priority areas

estimated time one month

Needs analysis in designated area, i.e. highest priority area identified by consultations — estimated time two months

Development of specifications and trial version of tests (S and L). S is to be assessed in controlled live interview only with specified interactions (cf. IELTS Speaking Test)

estimated time two months

Trials and development of interpretative procedures

estimated time one month

Review of trial tests, development of first live version, and development of interviewer training package

estimated time one month

Outcome:

One live version of the tests to be secure and centrally administered

Indicative Staffing:

Project Director) Testing Expert)	equivalent to 0.2 of a full-time position
Test Writers (applied linguists who are native or high proficiency speakers of Japanese)	2 full-time positions
Secretary (bilingual)	0.2 of a full-time position
Advisory Committee (largely native Japanese speaking teachers, applied linguists and teachers of Japanese plus an ASC representative)	

Indicative costing (estimated for 7 months)

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent to 0.2 of a full-time position)		$ 10,000
2 test-writers (estimated at middle of lecturer scale)		$ 58,500
Bilingual secretary		$ 4,500
TOTAL SALARY COSTS		$ 73,000
Advisory Committee costs for 2 meetings	estimated	$ 4,000
Research costs:		
- review of literature and courses and incidentals	estimated	$ 2,000
- consultations with user groups	estimated	$ 2,000
- needs analysis	estimated	$ 5,000
Development of interviewer training package	estimated	$ 6,000

Trials:		
- travel	estimated	$ 6,000
- computer costs: contributed by institution		-
Consumables	estimated	$ 3,000
Equipment	estimated	$ 5,000
Marketing (N.B. No salary costs included)	estimated	$ 5,000
Office space: contributed by institution		-
Institutional overheads	nominal	$ 15,000
TOTAL		$126,000

It is estimated that this test should become self-funding soon after release of the first live version of the test. Consequently no on-going development costs are included for additional parallel versions.

This test should be marketable in other countries that receive Japanese tourists (e.g. South East Asia, Europe, North America) and this may help to recover initial and on-going development costs.

Note: This project is related to the development of test instruments in the 0+ to 1 proficiency range (Project Proposal 3) and some saving may be effected if the two projects proceed together as one project.

PROJECT PROPOSAL 5

TEST BATTERY REQUIRED TO ASSESS CANDIDATES ON THE ASIAN LANGUAGE PROFICIENCY RATING SCALE (SPECIAL PURPOSE VERSION) IN THE AREAS OF JAPANESE FOR BUSINESS, LAW, SCIENCE, TECHNOLOGY OR ANY OTHER FIELD
(relating to Instruments 8 and 9)

Aim.

To develop a test battery able to assess the special purpose proficiency of candidates in Speaking, Listening, Reading and Writing in one of the areas of business, law, science, technology, etc.

This project is costed for one test battery only. Others may be developed on this model. The cost of subsequent versions will be less than for the initial version. This project pre-supposes development of the generic and special purpose scales.

Stages of the project (Tests in Japanese for Business Purposes):

Review of literature and courses in Japanese for Business, and consultation with industry and relevant institutions

estimated time one month

Development of specifications and trial version of tests (significant research required, including needs analysis); the form of the Speaking test will depend on the preferred form identified in developing Instrument 2 (i.e. live interview, recorded interview, etc.)

estimated time eight months

Trials and development of interpretative procedures

estimated time 1.5 months

Review of trial tests, development of first live version, and preparation of interviewer and administrator training package

estimated time 1.5 months

Outcomes:

One live version of the tests to be secure and centrally administered

Advice to other users on development of own tests

Additional items to be placed in test item banks

Indicative staffing:

Project Director) Testing Expert)	equivalent to 0.2 of a full-time position
Test Writers (applied linguists who are native or high proficienc / speakers of Japanese)	2 full-time positions
Secretary (bilingual)	0.2 of a full-time position
Advisory Committee (largely local teachers, applied linguists and user representatives plus an ASC representative)	

Indicative costing (estimated for 12 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent to 0.2 full-time position)		$ 16,300
2 test writers (estimated at middle of lecturers scale)		$100,000
Bilingual secretary		$ 7,000
TOTAL SALARY COSTS		$123,300
Advisory Committee costs for 3 meetings	estimated	$ 6,000
Research costs:		
- literature reviews and incidentals	estimated	$ 3,000
- consultations with user groups and institutions	estimated	$ 6,000
Trials:		
- travel	estimated	$ 6,000
- computer costs: contributed by . institution		-
Consumables	estimated	$ 3,000
Equipment	estimated	$ 5,000

Marketing	estimated	$ 3,000
Office space: contributed by institution		-
Institutional overheads	nominal	$ 17,000
TOTAL		172,300

It is estimated that the test should become self-funding, though how soon this will occur will depend on the number of candidates wishing to take the test and the effectiveness of the marketing strategies adopted in Australia and overseas. Some on-going development costs will need to be met if additional parallel versions of the test are to be available for reasons of security. The extent of the subsidy that may required until the test is self-funding is estimated at 50% of the development costs, making a total additional cost of $86,150.

Total development costs before the test becomes fully self-funding $258,450.

Note: This.project is closely related to the development of the generic scale (Project Proposal 1) and there will be substantial saving if the two development activities proceed together as a single project.

PROJECT PROPOSAL 6

DEVELOPMENT OF THE SPECIAL PURPOSE VERSION OF THE SCALE
(relating to Instrument 10)

Aim:

To develop a special purpose version of the Asian Language Proficiency Rating Scale as a basis for the assessment of proficiency in Japanese in special purpose contexts and as an overarching framework within which special purpose courses can be rationally and coherently developed.

The scale will have four sub-scales in Speaking, Listening, Writing and Reading.

It is recommended that this project proceed after development of the generic scale (the Asian Languages Proficiency Rating Scale (Japanese)). Some saving may be effected if it were to proceed in tandem with the development of Instruments 8 and 9 - the special purpose tests in Japanese for Special Purposes in the areas of business, law, science and technology (Project Proposal 5).

It should be noted that there is currently a project in progress at Griffith University to develop special purpose versions of the ASLPR. In this project significant theoretical issues are being worked through and there would be some advantage in delaying the commencement of this project until the ASLPR For Special Purposes project is nearing completion. In any case, since it is proposed to re-develop the ASLPR for Japanese as the Asian Languages Proficiency Rating Scale, there would seem to be advantage in awaiting the ASLPR for Special Purposes project so as to draw on the theoretical and practical lessons learned in that project.

Using business as the example, it is proposed to call the new scale the Asian Languages Proficiency Rating Scale for Special Purposes (Japanese for Business Purposes). Other fields would be substituted for "business" as different versions of the scale were developed.

The new scale will serve as a model for parallel scales in other Asian languages. It is also expected that it will serve as a ready example for other special purpose versions.

Stages of the Project:

Formation of a project advisory committee to include representatives of key users (including employers), experts in language testing (especially proficiency assessment), native speakers of Japanese, and highly proficient native English speaking teachers of Japanese.

Literature review in the area of special purpose assessment and rating scales but, in particular, of the work on the ASLPR for Special Purposes.

estimated time one month

Re-consideration of the generic scale (i.e. the Asian Languages Proficiency Rating Scale) in the context of special purpose assessment. Formulation of a tentative special purpose version.

estimated time one month

Face-to-face interviews with at least 50 learners in each of the immediate target fields (e.g. Business, Law, Science and Technology) to re-consider the Asian Languages Proficiency Rating Scale in special purpose contexts throughout the proficiency range in different courses, in different locations, and of different ages from adolescents up and development of supplementary descriptors in the special purpose contexts.

estimated time three months

Formal trialling and validation to assess the coherence and validity of the descriptors and the extent to which different raters are able to ascribe similar ratings to the same learners. This will include making videos of at least 15 or 16 learners in each target area throughout the proficiency range and having ratings assigned by 20 native and high proficiency non-native speakers of Japanese.

estimated time three months

Review and modification of the scale in the light of the trial results.

estimated time one month

Release of the Asian Languages Proficiency Rating Scale for Special Purposes (Japanese for Business Purposes, etc.).

Timelines:

Given that there is close interrelationship between this project and the project to develop Instruments 8 and 9, with input from one to the other, the likely time required is estimated at 9 months. However, if the two projects are developed as a single project with no additional staffing allocated, it would be reasonable to estimate that an additional time of 6 months over and above that indicated for the development of Instruments 8 and 9 would be required.

Indicative staffing:

Project Director) Testing Expert)	equivalent to 0.2 of a full-time position
Project Coordinator (an applied linguist who is native or high proficiency speaker of Japanese with an understanding of proficiency rating scales)	1 full-time position
Secretary (bilingual)	0.2 of a full-time position
Advisory Committee (see Project Proposal 1)	

Indicative costing (estimated for 9 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent 0.2 of a full-time position)		$ 12,200
Project Coordinator (estimated at middle of lecturers scale)		$ 37,500
Bilingual secretary		$ 5,200
TOTAL SALARY COSTS		$ 54,900
Advisory Committee costs for 3 meetings	estimated	$ 6,000
Research costs:		
- literature reviews and incidentals	estimated	$ 1,500
- consultations with user groups and institutions and informants	estimated	$ 4,000

Development interviews estimated cost including travel and informant time	estimated	$	6,000
Trials:			
- preparation of videos and other trial material	estimated	$	5,000
- travel	estimated	$	6,000
- computer costs: contributed by institution			-
Consumables	estimated	$	2,000
Equipment	estimated	$	2,000
Office space: contributed by institution			-
Institutional overheads	nominal	$	12,000
TOTAL		$	99,400

Funding sources:

In addition to Australian government sources, contributions may be available from Canada, Australian industry, the Japan Foundation and Japanese industry.

PROJECT PROPOSAL 7

TESTING IN THE SCHOOL CONTEXT: PROFICIENCY ASSESSMENT, DIAGNOSTIC TESTS AND ATTAINMENT TESTS

(relating to Instruments 11, 12 and 13)

Aims:

1. To develop proficiency scales for use in the school context, to provide an overarching framework within which curricula can be rationally and coherently developed, and to provide a basis on which to develop instruments to assess the proficiency of school learners.

2. To develop diagnostic and attainment tests for use with school learners especially where programs are based on the national curriculum for Japanese.

Management of the project:

Given the scope of the project and its components, the possibility of international cooperation should be pursued.

The Australian participants should be familiar with:

- the Japanese national curriculum project
- NAFLaSSL

The team should include:

- proficiency and testing experts
- native speakers of Japanese
- high proficiency expert teachers of Japanese (including native English speakers)
- education experts in primary and secondary child development specially related to language acquisition

Stages of the project:

1. Development of a proficiency continum for school students (i.e. second language learners)

2. Development of specifications for the following tests and instruments:

   a. instruments to relate learners to the proficiency scale

   b. diagnostic tests related to the proficiency scale and the national curriculum

   c. attainment tests related to the national curriculum

   d. formulation of proficiency profiles

3. Preparation of instruments including sample tests, exemplar tasks, item banks and guidelines for teacher made tests or other assessment procedures.

4. Development of a pre- and in-service teacher education package focussing on the development, use and interpretation of tests, recording of observations, etc. and interfacing with programs that address teachers' linguistic comprehension and methodology.

Possible applications:

1. A model for other languages
2. A stimulus for further curriculum development
3. Support for the process of national assessment of languages
4. Informative, consistent performance-based reporting for a wide variety of users (e.g. teachers, employers, parents, learners, etc.)
5. Basis for comparison of performance by learners passing through different programs in different locations
6. Probable international use
7. Facilitating student mobility
8. Facilitating transition from primary to secondary school and secondary to tertiary
9. Facilitating diagnosis, enhanced teaching effectiveness and articulation of student learning pathways.

Timelines

There will be a series of related and overlapping stages

Years 0_____1_____2_____3

Stage 1_____/

Stage 2 _____/

Stage 3 _____/

Stage 4 _____/

Note: These timelines may be contracted if resources are available to increase staffing.

Indicative costs:

Stage 1 (scale development and trialling)	$ 150,000
Stage 2 (development of specifications)	$ 50,000
Stage 3 (development and trialling of tests)	$ 300,000
Stage 4 (development of teacher education packages and commencement of teacher in-service education)	$ 300,000
TOTAL	$ 800,000

Note: This project would set up a model and provide data for similar projects in the other languages in which the ASC has interest.

PROJECT PROPOSAL 8

TESTS OF JAPANESE FOR ACADEMIC PURPOSES
(relating to Instrument 14)

Aim:

To develop speaking and writing tests to complement the Japanese Language Proficiency Test.

Stages of the project:

Review of the literature on Japanese for academic purposes and investigation of the work of the Society for the Teaching of Japanese, especially in the assessment of Speaking

estimated time 1.5 months

Development of specifications, trial version of tests, and interviewer training package. The form of the Speaking test will depend on the preferable form (e.g. live or recorded interview) identified in developing Instrument 2)

estimated time 4 months

Trials and development of interpretative procedures

estimated time 2 months

Review of trial tests and development of first live version

estimated time 1.5 months

Outcomes:

One live version of the tests to be secure and centrally administered

Additional items to be placed in test item banks.

Indicative staffing:

Project Director) equivalent to 0.2 of a full-time
Testing Expert) position

Test Writers (applied linguists who are native or high proficiency speakers of Japanese)	2 full-time positions
Secretary (bilingual)	0.2 of a full-time position
Advisory Committee (largely native Japanese speaking teachers, applied linguists and tertiary teachers of Japanese plus an ASC representative)	

Indicative costing (estimated for 9 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent 0.2 of a full-time position)		$ 12,000
2 test writers (estimated at middle of lecturers scale)		$ 70,000
Bilingual secretary		$ 6,000
TOTAL SALARY COSTS		$ 88,000
Advisory Committee costs for 3 meetings	estimated	$ 6,000
Research costs:		
- literature reviews and incidentals	estimated	$ 3,000
- consultations with user groups, especially institutions	estimated	$ 2,000
Development of interviewer training package	estimated	$ 6,000
Trials:		
- travel	estimated	$ 6,000
- computer costs: contributed by institution		-
Consumables	estimated	$ 3,000
Equipment	estimated	$ 5,000

Marketing (N.B. no salary time allocated)	estimated	$ 2,000
Office space: contributed by institution		-
Institutional overheads	nominal	$ 15,000
	TOTAL	$ 136,000

It is estimated that the test should become self-funding within a period of two years after release of the first live version of the test. Consequently some on-going development costs will need to be met if additional parallel versions of the test are to be available for reasons of security. The estimated cost of the on-going development activities is 50% of the development costs in the first year after release of the first live version of the test and 25% in the second year making a total additional cost of $102,000.

Total development costs before the test becomes fully self-funding $238,000.

Note: This project is closely related to the development of the generic scale (Project Proposal 1) and there will be substantial saving if the two development activities proceed together as a single project.

PROJECT PROPOSAL 9

TEST OF SOCIOCULTURAL AND SOCIOLINGUISTIC KNOWLEDGE IN JAPANESE
(relating to Instrument 15)

Aim:

To develop tests of sociocultural and sociolinguistic knowledge in Japanese for

(a) foreign language learners of Japanese
(b) non-speakers of Japanese

Stages of the project:

Review of the literature on cultural understanding and its assessment especially, though not exclusively, in relation to Japanese

estimated time two months

Development of specifications and trial versions of tests (NB. significant research into theoretical aspects of the sociocultural and sociolinguistic aspects of Japanese, their implications for second or foreign language learners of different ages and different proficiency levels, and their assessment with both learners of Japanese and non-learners will be required)

estimated time twelve months

Trials and development of interpretative procedures

estimated time two months

Review of trial tests and development of first live version

estimated time two months

Outcome:

One live version of each of two tests to be secure and centrally administered. One version will be for use with learners of Japanese as a foreign language and the other will be for use with non-learners/speakers of Japanese.

Indicative staffing:

Project Director) Tester)	equivalent to 0.5 of a full-time position (NB. One of these persons should be expert in the design and construction of tests and the other should be an expert in Japanese language and culture)
Test Writer (applied linguist who is a native or high proficiency speaker of Japanese)	1 full-time position
Secretary (bilingual)	0.5 of a full-time position
Advisory Committee (largely native Japanese speaking teachers, applied linguists and tertiary teachers of Japanese plus an ASC representative)	

Indicative costing (estimated for 18 months):

Salary costs include 25% for academics, 15% for test writers and secretary.

Director plus tester (equivalent 0.5 of a full-time position, costed at $60,000 average)		$ 56,250
1 test writer (estimated at middle of lecturers scale)		$ 75,000
Bilingual secretary		$ 25,950
TOTAL SALARY COSTS		$ 157,200
Advisory Committee costs for 4 meetings	estimated	$ 8,000
Research costs:		
- literature reviews and incidentals	estimated	$ 4,000
- consultations with user groups, especially institutions and industry	estimated	$ 4,000
Trials:		
- travel	estimated	$ 6,000

- computer costs: contributed by institution		-
Consumables	estimated	$ 6,000
Equipment	estimated	$ 5,000
Marketing	estimated	$ 3,000
Office space: contributed by institution		-
Institutional overheads	nominal	$ 20,000
TOTAL		$ 213,200

The test should be sold in order to re-coup some of the development costs and, in particular, to fund development of additional versions of the test (though this is less necessary with this test than it is with the language tests). The attempt should be made to market the tests overseas as well as in Australia.